GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2006, 23:00:48 ; Search time 18970 Seconds
(without alignments)
11686.318 Million cell updates/sec

Title: US-10-677-943A-23
Perfect score: 3900
Sequence: 1 atgaaggttgcaggaggact.....gaaaaaaaaaaaaaaaaaa 3900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3900	100.0	3900	6	AX427610	AX427610 Sequence
	2	3885	99.6	3885	6	CQ971706	CQ971706 Sequence
	3	3885	99.6	3885	8	AY054986	AY054986 Homo sapi
	4	3870.6	99.2	3885	8	AY154460	AY154460 Homo sapi
	5	3593.2	92.1	3926	6	AX704821	AX704821 Sequence
	6	3559.2	91.3	3830	6	AX704823	AX704823 Sequence
	7	3377.4	86.6	5859	6	AX459881	AX459881 Sequence
	8	3377.4	86.6	6939	6	AX459873	AX459873 Sequence
	9	3377.4	86.6	6939	6	AX459891	AX459891 Sequence
	10	3311.2	84.9	3489	6	AX478549	AX478549 Sequence
	11	2565.6	65.8	2753	6	CQ731113	CQ731113 Sequence
c	12	1595.6	40.9	157141	8	AC011470	AC011470 Homo sapi
	13	1595.6	40.9	193609	8	AC024580	AC024580 Homo sapi
	14	1478.2	37.9	3542	4	AY721594	AY721594 Bos tauru
	15	1301.8	33.4	3405	9	AY329487	AY329487 Mus muscu
	16	1300.8	33.4	3432	9	AY329484	AY329484 Mus muscu
	17	1297.6	33.3	3480	9	AY196361	AY196361 Mus muscu
	18	1296	33.2	3470	9	AY196362	AY196362 Mus muscu
	19	1294.4	33.2	3453	9	AY329486	AY329486 Mus muscu
	20	1292.8	33.1	3447	6	AX427592	AX427592 Sequence
	21	1292.8	33.1	3447	9	AF074018	AF074018 Mus muscu
	22	1292.8	33.1	3454	9	AY329490	AY329490 Mus muscu
	23	1291.2	33.1	3534	9	BC053384	BC053384 Mus muscu
	24	1158.8	29.7	3375	9	AY329489	AY329489 Mus muscu
	25	1155.6	29.6	3372	9	AY329485	AY329485 Mus muscu
	26	1149.6	29.5	1157	6	AX427588	AX427588 Sequence
	27	1073.4	27.5	1075	6	AX427590	AX427590 Sequence
	28	1025	26.3	3303	9	AY329488	AY329488 Mus muscu
	29	1021.8	26.2	3303	9	AY329491	AY329491 Mus muscu
c	30	988.6	25.3	167509	14	AC012107	AC012107 Homo sapi
	31	856	21.9	196031	9	AC127352	AC127352 Mus muscu
	32	856	21.9	196193	14	AC155171	AC155171 Mus muscu
	33	856	21.9	205087	14	AC148980	AC148980 Mus muscu
	34	727.4	18.7	61925	14	AC164677	AC164677 Bos tauru
	35	679.2	17.4	3282	8	AY154469	AY154469 Homo sapi
c	36	647.8	16.6	231371	14	AC132633	AC132633 Rattus no
	37	647.8	16.6	323826	14	AC110429	AC110429 Rattus no
	38	643.6	16.5	1576	9	F143559S07	AF143565 Mus muscu
	39	640.4	16.4	203222	14	AC159288	AC159288 Mus muscu
	40	528.4	13.5	167509	14	AC012107	AC012107 Homo sapi
	41	427	10.9	3281	9	AY673647	AY673647 Mus muscu
	42	425.6	10.9	2964	9	AY596199	AY596199 Mus muscu
	43	422.4	10.8	3190	6	AX714395	AX714395 Sequence
	44	422.4	10.8	3190	8	AK056688	AK056688 Homo sapi
	45	422.4	10.8	3333	8	AF442488	AF442488 Homo sapi

ALIGNMENTS

RESULT 1
AX427610
LOCUS AX427610 3900 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 23 from Patent WO0232955.
ACCESSION AX427610
VERSION AX427610.1 GI:21537730
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Nelson,L.M. and Tong,Z.-B.
TITLE Human gene critical to fertility
JOURNAL Patent: WO 0232955-A 23 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
source 1. .3900
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 1. .3603
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35973.1"
/db_xref="GI:21537731"
/translation="MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNL
SSQPCIKMEGDKSLTFSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIE
NANVECLALLLHEYYGASLAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTD
QGPSKEKVPGISQAVQQDSATAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGA
TAAETEEQGHGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWG
FRPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTE
FISREWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIR
SLLRKVLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQ
KTQGLRAIINNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFH
QLTPRGVVRRCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFH
MNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQA
GFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPG
DTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDV
KGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAM

KTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRM
ACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLSLAGNKVTDQGVTPL
SDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSNNSLGNEGVNLLC
RSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPVEDNGVKLLCE
VMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDGGVAALCEG
LKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMMKLCSAF
ACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN"

ORIGIN

Query Match 100.0%; Score 3900; DB 6; Length 3900;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3900; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAAGGTTGCAGGAGGACTTGAACTTGGAGCTGCTGCTCTGCTCTCAGCATCACCACGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGGTTGCAGGAGGACTTGAACTTGGAGCTGCTGCTCTGCTCTCAGCATCACCACGT 60

Qy         61 GCTCTTGTCACTCTTTCCACAGGTCCTACTTGCTCTATATTACCAAAGAATCCACTTTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCTTGTCACTCTTTCCACAGGTCCTACTTGCTCTATATTACCAAAGAATCCACTTTTC 120

Qy        121 CCCCAAAACCTGAGCTCTCAGCCTTGTATCAAGATGGAAGGAGACAAATCGCTCACCTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCCAAAACCTGAGCTCTCAGCCTTGTATCAAGATGGAAGGAGACAAATCGCTCACCTTT 180

Qy        181 TCCAGCTACGGGCTGCAATGGTGTCTCTATGAGCTAGACAAGGAAGAATTTCAGACATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCAGCTACGGGCTGCAATGGTGTCTCTATGAGCTAGACAAGGAAGAATTTCAGACATTC 240

Qy        241 AAGGAATTACTAAAGAAGAAATCTTCAGAATCGACCACATGCTCTATTCCACAGTTTGAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGAATTACTAAAGAAGAAATCTTCAGAATCGACCACATGCTCTATTCCACAGTTTGAA 300

Qy        301 ATCGAGAATGCCAACGTGGAATGTCTGGCACTCCTCTTGCATGAGTATTATGGAGCATCG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCGAGAATGCCAACGTGGAATGTCTGGCACTCCTCTTGCATGAGTATTATGGAGCATCG 360

Qy        361 CTGGCCTGGGCTACGTCCATTAGCATCTTTGAAAACATGAACCTGCGAACCCTCTCGGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGCCTGGGCTACGTCCATTAGCATCTTTGAAAACATGAACCTGCGAACCCTCTCGGAG 420

Qy        421 AAGGCACGGGATGACATGAAAAGACATTCACCAGAAGATCCTGAAGCAACGATGACTGAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGGCACGGGATGACATGAAAAGACATTCACCAGAAGATCCTGAAGCAACGATGACTGAC 480

Qy        481 CAAGGACCAAGCAAGGAAAAAGTGCCAGGAATTTCACAAGCTGTGCAACAAGATAGTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAAGGACCAAGCAAGGAAAAAGTGCCAGGAATTTCACAAGCTGTGCAACAAGATAGTGCC 540

Qy        541 ACAGCTGCAGAGACAAAAGAACAGGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAGCTGCAGAGACAAAAGAACAGGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACA 600

Qy        601 GCAGCAGAGACAGAAGAACAAGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGCAGAGACAGAAGAACAAGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCA 660

Qy        661 GCAGAGACAGAAGAACAAGGACATGGAGGTGACACATGGGACTACAAGAGTCACGTGATG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCAGAGACAGAAGAACAAGGACATGGAGGTGACACATGGGACTACAAGAGTCACGTGATG 720

Qy        721 ACCAAATTCGCTGAGGAGGAGGATGTACGTCGTAGTTTTGAAAACACTGCTGCTGACTGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACCAAATTCGCTGAGGAGGAGGATGTACGTCGTAGTTTTGAAAACACTGCTGCTGACTGG 780

Qy        781 CCGGAAATGCAAACGTTGGCTGGTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCGGAAATGCAAACGTTGGCTGGTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGC 840

Qy        841 ACGGTGGTTCTGCACGGAAAGTCAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACGGTGGTTCTGCACGGAAAGTCAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTG 900

Qy        901 CTGTGCTGGGCGCAAGGTGGACTCTACCAGGGAATGTTCTCCTACGTCTTCTTCCTCCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTGTGCTGGGCGCAAGGTGGACTCTACCAGGGAATGTTCTCCTACGTCTTCTTCCTCCCC 960

Qy        961 GTTAGAGAGATGCAGCGGAAGAAGGAGAGCAGTGTCACAGAGTTCATCTCCAGGGAGTGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTAGAGAGATGCAGCGGAAGAAGGAGAGCAGTGTCACAGAGTTCATCTCCAGGGAGTGG 1020

Qy       1021 CCAGACTCCCAGGCTCCGGTGACGGAGATCATGTCCCGACCAGAAAGGCTGTTGTTCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCAGACTCCCAGGCTCCGGTGACGGAGATCATGTCCCGACCAGAAAGGCTGTTGTTCATC 1080

Qy       1081 ATTGACGGTTTCGATGACCTGGGCTCTGTCCTCAACAATGACACAAAGCTCTGCAAAGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATTGACGGTTTCGATGACCTGGGCTCTGTCCTCAACAATGACACAAAGCTCTGCAAAGAC 1140

Qy       1141 TGGGCTGAGAAGCAGCCTCCGTTCACCCTCATACGCAGTCTGCTGAGGAAGGTCCTGCTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGGGCTGAGAAGCAGCCTCCGTTCACCCTCATACGCAGTCTGCTGAGGAAGGTCCTGCTC 1200

Qy       1201 CCTGAGTCCTTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCTGAGTCCTTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAG 1260

Qy       1261 GTCGTGTCTCCCCGTTACCTGTTAGTTAGAGGAATCTCCGGGGGAACAAAGAATCCACTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTCGTGTCTCCCCGTTACCTGTTAGTTAGAGGAATCTCCGGGGGAACAAAGAATCCACTTG 1320

Qy       1321 CTCCTTGAGCGCGGGATTGGTGAGCATCAGAAGACACAAGGGTTGCGTGCGATCATCAAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCCTTGAGCGCGGGATTGGTGAGCATCAGAAGACACAAGGGTTGCGTGCGATCATCAAC 1380

Qy       1381 AACCGTGAGCTGCTCGACCAGTGCCAGGTGCCCGCCGTGGGCTCTCTCATCTGCGTGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AACCGTGAGCTGCTCGACCAGTGCCAGGTGCCCGCCGTGGGCTCTCTCATCTGCGTGGCC 1440

Qy       1441 CTGCAGCTGCAGGACGTGGTGGGGGAGAGCGTCGCCCCCTTCAACCAAACGCTCACAGGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCAGCTGCAGGACGTGGTGGGGGAGAGCGTCGCCCCCTTCAACCAAACGCTCACAGGC 1500

Qy       1501 CTGCACGCCGCTTTTGCGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGCACGCCGCTTTTGCGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTC 1560

Qy       1561 AATCTGGAGGAAAGAGTTGTCCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AATCTGGAGGAAAGAGTTGTCCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGG 1620

Qy       1621 AATAGGAAGTCAGTGTTTGATGGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AATAGGAAGTCAGTGTTTGATGGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAG 1680

Qy       1681 CTCCGTGCTCTGTTTCACATGAACATCCTTCTCCCAGACAGCCACTGTGAGGAGTACTAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTCCGTGCTCTGTTTCACATGAACATCCTTCTCCCAGACAGCCACTGTGAGGAGTACTAC 1740

Qy       1741 ACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCGCCTTGTACTACGTGTTAGAGGGC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCGCCTTGTACTACGTGTTAGAGGGC 1800

Qy       1801 CTGGAAATCGAGCCAGCTCTCTGCCCTCTGTACGTTGAGAAGACAAAGAGGTCCATGGAG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGAAATCGAGCCAGCTCTCTGCCCTCTGTACGTTGAGAAGACAAAGAGGTCCATGGAG 1860

Qy       1861 CTTAAACAGGCAGGCTTCCATATCCACTCGCTTTGGATGAAGCGTTTCTTGTTTGGCCTC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTTAAACAGGCAGGCTTCCATATCCACTCGCTTTGGATGAAGCGTTTCTTGTTTGGCCTC 1920

Qy       1921 GTGAGCGAAGACGTAAGGAGGCCACTGGAGGTCCTGCTGGGCTGTCCCGTTCCCCTGGGG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTGAGCGAAGACGTAAGGAGGCCACTGGAGGTCCTGCTGGGCTGTCCCGTTCCCCTGGGG 1980

Qy       1981 GTGAAGCAGAAGCTTCTGCACTGGGTCTCTCTGTTGGGTCAGCAGCCTAATGCCACCACC 2040
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2006, 22:51:04 ; Search time 1455 Seconds
(without alignments)
17864.125 Million cell updates/sec

Title: US-10-677-943A-23
Perfect score: 3900
Sequence: 1 atgaaggttgcaggaggact.....gaaaaaaaaaaaaaaaaaa 3900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3900	100.0	3900	6	ABK48628	Abk48628 Human MAT
2	3900	100.0	3900	8	AAD49018	Aad49018 Human MAT
3	3885	99.6	3885	14	ADV04680	Adv04680 NALP5 cod
4	3593.2	92.1	3926	8	ADA45218	Ada45218 Human MAT
5	3559.2	91.3	3830	8	ADA45220	Ada45220 Human MAT
6	3377.4	86.6	5859	6	AAL47135	Aal47135 Pyrin dom
7	3377.4	86.6	6939	6	AAL47131	Aal47131 Pyrin dom
8	3377.4	86.6	6939	6	AAL47140	Aal47140 Pyrin dom
9	3311.2	84.9	3489	6	AAD41224	Aad41224 Human EMB
10	3166.2	81.2	4035	6	AAL44356	Aal44356 Human PYR
11	3166.2	81.2	4035	10	ADF94744	Adf94744 Human PYR
12	2965.2	76.0	3226	6	ABX97181	Abx97181 Human NOV
13	1903.6	48.8	2820	12	ADN62078	Adn62078 Human cDN
14	1292.8	33.1	3447	6	ABK48611	Abk48611 Mouse MAT
15	1292.8	33.1	3447	8	AAD49001	Aad49001 Mouse MAT
16	1149.6	29.5	1157	6	ABK48609	Abk48609 Human MAT
17	1149.6	29.5	1157	8	AAD48999	Aad48999 Human MAT
18	1073.4	27.5	1075	6	ABK48610	Abk48610 Human MAT
19	1073.4	27.5	1075	8	AAD49000	Aad49000 Human MAT
20	679.2	17.4	3521	14	ADV44739	Adv44739 Human nuc
21	480.8	12.3	2808	12	ADP47735	Adp47735 Human inf
22	475.8	12.2	3039	12	ADP47733	Adp47733 Human inf
23	424	10.9	2985	6	ADE36408	Ade36408 Human PAA
24	424	10.9	2985	12	ADJ19330	Adj19330 Human PAN
25	422.4	10.8	2982	10	ADF94764	Adf94764 Human PYR
26	422.4	10.8	3190	10	ADA53511	Ada53511 Human cod
27	422.4	10.8	3368	6	AAL44366	Aal44366 Human PYR
28	422.4	10.8	3368	10	ADF94760	Adf94760 Human PYR
29	419.2	10.7	2575	6	AAI70683	Aai70683 Human nuc
30	419.2	10.7	2767	8	ABT16018	Abt16018 NOVX rela
31	419.2	10.7	3172	6	ABN99366	Abn99366 Human sec
32	419.2	10.7	3310	13	ACN40690	Acn40690 Tumour-as
33	418.8	10.7	2883	12	ADO41637	Ado41637 Novel hum
34	405.8	10.4	3186	6	AAL44363	Aal44363 Human PYR
35	405.8	10.4	3186	10	ADF94754	Adf94754 Human PYR
36	401.2	10.3	3108	6	ADE36416	Ade36416 Human PAA
37	401.2	10.3	3108	12	ADJ19338	Adj19338 Human PAN
38	401.2	10.3	3218	12	ADJ19382	Adj19382 Human PAN
39	401.2	10.3	3300	6	AAL47129	Aal47129 Pyrin dom
40	401.2	10.3	3731	12	ADP47723	Adp47723 Human Mon
41	394	10.1	2799	12	ADH18983	Adh18983 Human cel
42	387.8	9.9	3531	6	AAL47128	Aal47128 Pyrin dom
43	387.4	9.9	3540	14	ADV44744	Adv44744 Human nuc
44	386.2	9.9	3189	5	AAI67185	Aai67185 Nucleotid
45	386.2	9.9	4944	10	ADF14464	Adf14464 Human end

ALIGNMENTS

RESULT 1
ABK48628
ID ABK48628 standard; cDNA; 3900 BP.
XX
AC ABK48628;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER cDNA.
XX
KW Human; gene; ss; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .3603
FT /*tag= a
FT /product= "Human MATER"
XX
PN WO200232955-A1.
XX
PD 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
DR WPI; 2002-454595/48.
DR P-PSDB; AAU79526.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 11; Page 80-85; 93pp; English.

XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,
CC the MATER protein is useful as a contraceptive agent, or for influencing
CC (either inhibiting or enhancing) fertility and can be used to detect a
CC predisposition to infertility or reduced fertility, or for presymptomatic
CC screening of an individual for infertility/reduced fertility. The protein
CC and polynucleotide are also useful for detecting an excess or deficiency,
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC human or a mouse) or for screening for a compound useful in influencing
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA.
CC Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to
CC published human genomic DNA sequences, while the remainder of the
CC sequence was determined by direct cloning of human ovarian cDNAs
XX
SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;

Query Match 100.0%; Score 3900; DB 6; Length 3900;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3900; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAAGGTTGCAGGAGGACTTGAACTTGGAGCTGCTGCTCTGCTCTCAGCATCACCACGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGGTTGCAGGAGGACTTGAACTTGGAGCTGCTGCTCTGCTCTCAGCATCACCACGT 60

Qy         61 GCTCTTGTCACTCTTTCCACAGGTCCTACTTGCTCTATATTACCAAAGAATCCACTTTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCTTGTCACTCTTTCCACAGGTCCTACTTGCTCTATATTACCAAAGAATCCACTTTTC 120

Qy        121 CCCCAAAACCTGAGCTCTCAGCCTTGTATCAAGATGGAAGGAGACAAATCGCTCACCTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCCAAAACCTGAGCTCTCAGCCTTGTATCAAGATGGAAGGAGACAAATCGCTCACCTTT 180

Qy        181 TCCAGCTACGGGCTGCAATGGTGTCTCTATGAGCTAGACAAGGAAGAATTTCAGACATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCAGCTACGGGCTGCAATGGTGTCTCTATGAGCTAGACAAGGAAGAATTTCAGACATTC 240

Qy        241 AAGGAATTACTAAAGAAGAAATCTTCAGAATCGACCACATGCTCTATTCCACAGTTTGAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGAATTACTAAAGAAGAAATCTTCAGAATCGACCACATGCTCTATTCCACAGTTTGAA 300

Qy        301 ATCGAGAATGCCAACGTGGAATGTCTGGCACTCCTCTTGCATGAGTATTATGGAGCATCG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCGAGAATGCCAACGTGGAATGTCTGGCACTCCTCTTGCATGAGTATTATGGAGCATCG 360

Qy        361 CTGGCCTGGGCTACGTCCATTAGCATCTTTGAAAACATGAACCTGCGAACCCTCTCGGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGCCTGGGCTACGTCCATTAGCATCTTTGAAAACATGAACCTGCGAACCCTCTCGGAG 420

Qy        421 AAGGCACGGGATGACATGAAAAGACATTCACCAGAAGATCCTGAAGCAACGATGACTGAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGGCACGGGATGACATGAAAAGACATTCACCAGAAGATCCTGAAGCAACGATGACTGAC 480

Qy        481 CAAGGACCAAGCAAGGAAAAAGTGCCAGGAATTTCACAAGCTGTGCAACAAGATAGTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAAGGACCAAGCAAGGAAAAAGTGCCAGGAATTTCACAAGCTGTGCAACAAGATAGTGCC 540

Qy        541 ACAGCTGCAGAGACAAAAGAACAGGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAGCTGCAGAGACAAAAGAACAGGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACA 600

Qy        601 GCAGCAGAGACAGAAGAACAAGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGCAGAGACAGAAGAACAAGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCA 660

Qy        661 GCAGAGACAGAAGAACAAGGACATGGAGGTGACACATGGGACTACAAGAGTCACGTGATG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCAGAGACAGAAGAACAAGGACATGGAGGTGACACATGGGACTACAAGAGTCACGTGATG 720

Qy        721 ACCAAATTCGCTGAGGAGGAGGATGTACGTCGTAGTTTTGAAAACACTGCTGCTGACTGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACCAAATTCGCTGAGGAGGAGGATGTACGTCGTAGTTTTGAAAACACTGCTGCTGACTGG 780

Qy        781 CCGGAAATGCAAACGTTGGCTGGTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCGGAAATGCAAACGTTGGCTGGTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGC 840

Qy        841 ACGGTGGTTCTGCACGGAAAGTCAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACGGTGGTTCTGCACGGAAAGTCAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTG 900

Qy        901 CTGTGCTGGGCGCAAGGTGGACTCTACCAGGGAATGTTCTCCTACGTCTTCTTCCTCCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTGTGCTGGGCGCAAGGTGGACTCTACCAGGGAATGTTCTCCTACGTCTTCTTCCTCCCC 960

Qy        961 GTTAGAGAGATGCAGCGGAAGAAGGAGAGCAGTGTCACAGAGTTCATCTCCAGGGAGTGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTAGAGAGATGCAGCGGAAGAAGGAGAGCAGTGTCACAGAGTTCATCTCCAGGGAGTGG 1020

Qy       1021 CCAGACTCCCAGGCTCCGGTGACGGAGATCATGTCCCGACCAGAAAGGCTGTTGTTCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCAGACTCCCAGGCTCCGGTGACGGAGATCATGTCCCGACCAGAAAGGCTGTTGTTCATC 1080

Qy       1081 ATTGACGGTTTCGATGACCTGGGCTCTGTCCTCAACAATGACACAAAGCTCTGCAAAGAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATTGACGGTTTCGATGACCTGGGCTCTGTCCTCAACAATGACACAAAGCTCTGCAAAGAC 1140

Qy       1141 TGGGCTGAGAAGCAGCCTCCGTTCACCCTCATACGCAGTCTGCTGAGGAAGGTCCTGCTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGGGCTGAGAAGCAGCCTCCGTTCACCCTCATACGCAGTCTGCTGAGGAAGGTCCTGCTC 1200

Qy       1201 CCTGAGTCCTTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCTGAGTCCTTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAG 1260

Qy       1261 GTCGTGTCTCCCCGTTACCTGTTAGTTAGAGGAATCTCCGGGGAACAAAGAATCCACTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTCGTGTCTCCCCGTTACCTGTTAGTTAGAGGAATCTCCGGGGAACAAAGAATCCACTTG 1320

Qy       1321 CTCCTTGAGCGCGGGATTGGTGAGCATCAGAAGACACAAGGGTTGCGTGCGATCATCAAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCCTTGAGCGCGGGATTGGTGAGCATCAGAAGACACAAGGGTTGCGTGCGATCATCAAC 1380

Qy       1381 AACCGTGAGCTGCTCGACCAGTGCCAGGTGCCCGCCGTGGGCTCTCTCATCTGCGTGGCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AACCGTGAGCTGCTCGACCAGTGCCAGGTGCCCGCCGTGGGCTCTCTCATCTGCGTGGCC 1440

Qy       1441 CTGCAGCTGCAGGACGTGGTGGGGGAGAGCGTCGCCCCCTTCAACCAAACGCTCACAGGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCAGCTGCAGGACGTGGTGGGGGAGAGCGTCGCCCCCTTCAACCAAACGCTCACAGGC 1500

Qy       1501 CTGCACGCCGCTTTTGCGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGCACGCCGCTTTTGCGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTC 1560

Qy       1561 AATCTGGAGGAAAGAGTTGTCCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AATCTGGAGGAAAGAGTTGTCCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGG 1620

Qy       1621 AATAGGAAGTCAGTGTTTGATGGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AATAGGAAGTCAGTGTTTGATGGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAG 1680

Qy       1681 CTCCGTGCTCTGTTTCACATGAACATCCTTCTCCCAGACAGCCACTGTGAGGAGTACTAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTCCGTGCTCTGTTTCACATGAACATCCTTCTCCCAGACAGCCACTGTGAGGAGTACTAC 1740

Qy       1741 ACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCGCCTTGTACTACGTGTTAGAGGGC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2006, 23:16:05 ; Search time 9581 Seconds
(without alignments)
19044.945 Million cell updates/sec

Title: US-10-677-943A-23
Perfect score: 3900
Sequence: 1 atgaaggttgcaggaggact.....gaaaaaaaaaaaaaaaaaa 3900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3484.8	89.4	6435	11	DQ034508	DQ034508 Homo sapi
	2	2532.8	64.9	6435	11	DQ034509	DQ034509 Pan trogl
	3	1275.6	32.7	3475	4	AK087774	AK087774 Mus muscu
	4	720.6	18.5	791	7	CK000513	CK000513 AGENCOURT
c	5	671.4	17.2	684	5	BU630481	BU630481 UI-H-FL0-
c	6	668.6	17.1	693	5	BU634350	BU634350 UI-H-FL1-
c	7	627.8	16.1	642	5	BU618831	BU618831 UI-H-FH1-
	8	436.6	11.2	688	7	CR853520	CR853520 DKFZp469O
	9	424.4	10.9	624	7	CR853537	CR853537 DKFZp469M
	10	422.4	10.8	3359	4	BC021272	BC021272 Homo sapi
	11	422.4	10.8	3360	4	BC012789	BC012789 Homo sapi
	12	404.4	10.4	814	6	CB228956	CB228956 AGENCOURT
	13	379.8	9.7	776	7	CO814088	CO814088 AGENCOURT
	14	374.6	9.6	1913	4	AK016782	AK016782 Mus muscu
	15	372.8	9.6	808	7	CO804754	CO804754 AGENCOURT
	16	364	9.3	2814	11	DQ034514	DQ034514 Homo sapi
	17	347.6	8.9	786	7	CO802335	CO802335 AGENCOURT
	18	337.6	8.7	792	7	CO382844	CO382844 AGENCOURT
	19	337.2	8.6	2568	11	DQ052279	DQ052279 Homo sapi
	20	323.8	8.3	3054	11	DQ045972	DQ045972 Homo sapi
	21	323.8	8.3	3061	11	DQ045971	DQ045971 Pan trogl
	22	323.4	8.3	3018	11	DQ040143	DQ040143 Homo sapi
	23	322.4	8.3	3189	11	DQ045970	DQ045970 Homo sapi
	24	321.8	8.3	1119	3	BM454123	BM454123 AGENCOURT
	25	318.4	8.2	739	7	CO815646	CO815646 AGENCOURT
	26	315.2	8.1	2814	11	DQ034515	DQ034515 Pan trogl
	27	301.4	7.7	671	2	BB555431	BB555431 BB555431
	28	290.4	7.4	3018	11	DQ040144	DQ040144 Pan trogl
	29	284	7.3	817	7	CO797627	CO797627 AGENCOURT
	30	283.2	7.3	701	7	CO801395	CO801395 AGENCOURT
	31	275	7.1	3577	4	AK087843	AK087843 Mus muscu
	32	274.2	7.0	666	1	AV367637	AV367637 AV367637
	33	265.6	6.8	3218	4	AK054378	AK054378 Mus muscu
c	34	265.4	6.8	579	3	BM228619	BM228619 K0264H01-
c	35	260	6.7	565	3	BM226108	BM226108 K0222C06-
	36	244.8	6.3	3054	11	DQ045973	DQ045973 Pan trogl
c	37	239.4	6.1	515	3	BM227562	BM227562 K0245E10-
c	38	239	6.1	511	3	BM230035	BM230035 K0288H05-
c	39	239	6.1	512	3	BM226240	BM226240 K0225A02-
c	40	237.8	6.1	513	3	BM227869	BM227869 K0250F09-
c	41	236.6	6.1	519	2	BG071729	BG071729 H3102E07-
	42	236.6	6.1	519	2	BG084574	BG084574 H3102E07-
	43	236.6	6.1	589	6	CA559886	CA559886 K0264H01-
	44	233.4	6.0	743	7	CK635847	CK635847 UI-M-HN0-
	45	232.6	6.0	552	7	CN724433	CN724433 E0879B11-

ALIGNMENTS

RESULT 1
DQ034508
LOCUS DQ034508 6435 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens MATER gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION DQ034508
VERSION DQ034508.1 GI:66885717
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6435)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 6435)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.
FEATURES Location/Qualifiers
source 1. .6435
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
gene <1. .>6435
/gene="MATER"
/locus_tag="HC3997"
ORIGIN

Query Match 89.4%; Score 3484.8; DB 11; Length 6435;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3489; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
Qy        83 GTCCTACTTGCTCTATATTACCAAAGAATCCACTTTTCCCCCAAAACCTGAGCTCTCAGC 142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2915 GTCCTACTTGCTCTATATTACCAAAGAATCCACTTTTCCCCCAAAACCTGAGCTCTCAGC 2974

Qy       143 CTTGTATCAAGATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGT 202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2975 CTTGTATCAAGATGGAAGGAGACAAATCGCTCACCTTTTCCAGCTACGGGCTGCAATGGT 3034

Qy       203 GTCTCTATGAGCTAGACAAGGAAGAATTTCAGACATTCAAGGAATTACTAAAGAAGAAAT 262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3035 GTCTCTATGAGCTAGACAAGGAAGAATTTCAGACATTCAAGGAATTACTAAAGAAGAAAT 3094

Qy       263 CTTCAGAATCGACCACATGCTCTATTCCACAGTTTGAAATCGAGAATGCCAACGTGGAAT 322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3095 CTTCAGAATCGACCACATGCTCTATTCCACAGTTTGAAATCGAGAATGCCAACGTGGAAT 3154

Qy       323 GTCTGGCACTCCTCTTGCATGAGTATTATGGAGCATCGCTGGCCTGGGCTACGTCCATTA 382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3155 GTCTGGCACTCCTCTTGCATGAGTATTATGGAGCATCGCTGGCCTGGGCTACGTCCATTA 3214

Qy       383 GCATCTTTGAAAACATGAACCTGCGAACCCTCTCGGAGAAGGCACGGGATGACATGAAAA 442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3215 GCATCTTTGAAAACATGAACCTGCGAACCCTCTCGGAGAAGGCACGGGATGACATGAAAA 3274

Qy       443 GACATTCACCAGAAGATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAG 502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3275 GACATTCACCAGAAGATCCTGAAGCAACGATGACTGACCAAGGACCAAGCAAGGAAAAAG 3334

Qy       503 TGCCAGGAATTTCACAAGCTGTGCAACAAGATAGTGCCACAGCTGCAGAGACAAAAGAAC 562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3335 TGCCAGGAATTTCACAAGCTGTGCAACAAGATAGTGCCACAGCTGCAGAGACAAAAGAAC 3394

Qy       563 AGGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGAGACAGAAGAACAAG 622
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3395 AAGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGAGACAGAAGAACAAG 3454

Qy       623 AAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGAGACAGAAGAACAAGGAC 682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3455 AAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCAGCAGAGACAGAAGAACAAGNNN 3514

Qy       683 ATGGAGGTGACACATGGGACTACAAGAGTCACGTGATGACCAAATTCGCTGAGGAGGAGG 742
                                         ||||||||||||||||||||||||||||||||||||
Db      3515 NNNNNNNNNNNNNNNNNNNNNNCAAGAGTCACGTGATGACCAAATTCGCTGAGGAGGAGG 3574

Qy       743 ATGTACGTCGTAGTTTTGAAAACACTGCTGCTGACTGGCCGGAAATGCAAACGTTGGCTG 802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3575 ATGTACGTCGTAGTTTTGAAAACACTGCTGCTGACTGGCCGGAAATGCAAACGTTGGCTG 3634

Qy       803 GTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGCACGGTGGTTCTGCACGGAAAGT 862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3635 GTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGCACGGTGGTTCTGCACGGAAAGT 3694

Qy       863 CAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTGCTGTGCTGGGCGCAAGGTGGAC 922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3695 CAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTGCTGTGCTGGGCGCAAGGTGGAC 3754

Qy       923 TCTACCAGGGAATGTTCTCCTACGTCTTCTTCCTCCCCGTTAGAGAGATGCAGCGGAAGA 982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3755 TCTACCAGGGAATGTTCTCCTACGTCTTCTTCCTCCCCGTTAGAGAGATGCAGCGGAAGA 3814

Qy       983 AGGAGAGCAGTGTCACAGAGTTCATCTCCAGGGAGTGGCCAGACTCCCAGGCTCCGGTGA 1042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3815 AGGAGAGCAGTGTCACAGAGTTCATCTCCAGGGAGTGGCCAGACTCCCAGGCTCCGGTGA 3874

Qy      1043 CGGAGATCATGTCCCGACCAGAAAGGCTGTTGTTCATCATTGACGGTTTCGATGACCTGG 1102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3875 CGGAGATCATGTCCCGACCAGAAAGGCTGTTGTTCATCATTGACGGTTTCGATGACCTGG 3934

Qy      1103 GCTCTGTCCTCAACAATGACACAAAGCTCTGCAAAGACTGGGCTGAGAAGCAGCCTCCGT 1162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3935 GCTCTGTCCTCAACAATGACACAAAGCTCTGCAAAGACTGGGCTGAGAAGCAGCCTCCGT 3994

Qy      1163 TCACCCTCATACGCAGTCTGCTGAGGAAGGTCCTGCTCCCTGAGTCCTTCCTGATCGTCA 1222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3995 TCACCCTCATACGCAGTCTGCTGAGGAAGGTCCTGCTCCCTGAGTCCTTCCTGATCGTCA 4054

Qy      1223 CCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTGTCTCCCCGTTACCTGT 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4055 CCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTGTCTCCCCGTTACCTGT 4114

Qy      1283 TAGTTAGAGGAATCTCCGGGGAACAAAGAATCCACTTGCTCCTTGAGCGCGGGATTGGTG 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4115 TAGTTAGAGGAATCTCCGGGGAACAAAGAATCCACTTGCTCCTTGAGCGCGGGATTGGTG 4174

Qy      1343 AGCATCAGAAGACACAAGGGTTGCGTGCGATCATCAACAACCGTGAGCTGCTCGACCAGT 1402
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      4175 AGCATCAGAAGACACAAGGGTTGCGTGCGATCATGAACAACCGTGAGCTGCTCGACCAGT 4234

Qy      1403 GCCAGGTGCCCGCCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGG 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4235 GCCAGGTGCCCGCCGTGGGCTCTCTCATCTGCGTGGCCCTGCAGCTGCAGGACGTGGTGG 4294

Qy      1463 GGGAGAGCGTCGCCCCCTTCAACCAAACGCTCACAGGCCTGCACGCCGCTTTTGCGTTTC 1522
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db      4295 GGGAGAGCGTCGCCCCCTTCAACCAAACGCTCACAGGCCTGCACGCCGCTTTTGTGTTTC 4354

Qy      1523 ATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCC 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4355 ATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGGAAAGAGTTGTCC 4414

Qy      1583 TGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGGAATAGGAAGTCAGTGTTTGATG 1642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4415 TGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGGAATAGGAAGTCAGTGTTTGATG 4474

Qy      1643 GTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCACATGA 1702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4475 GTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAGCTCCGTGCTCTGTTTCACATGA 4534

Qy      1703 ACATCCTTCTCCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCC 1762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4535 ACATCCTTCTCCCAGACAGCCACTGTGAGGAGTACTACACCTTCTTCCACCTCAGTCTCC 4594

Qy      1763 AGGACTTCTGTGCCGCCTTGTACTACGTGTTAGAGGGCCTGGAAATCGAGCCAGCTCTCT 1822
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4595 AGGACTTCTGTGCCGCCTTGTACTACGTGTTAGAGGGCCTGGAAATCGAGCCAGCTCTCT 4654

Qy      1823 GCCCTCTGTACGTTGAGAAGACAAAGAGGTCCATGGAGCTTAAACAGGCAGGCTTCCATA 1882
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4655 GCCCTCTGTACGTTGAGAAGACAAAGAGGTCCATGGAGCTTAAACAGGCAGGCTTCCATA 4714

Qy      1883 TCCACTCGCTTTGGATGAAGCGTTTCTTGTTTGGCCTCGTGAGCGAAGACGTAAGGAGGC 1942
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4715 TCCACTCGCTTTGGATGAAGCGTTTCTTGTTTGGCCTCGTGAGCGAAGACGTAAGGAGGC 4774

Qy      1943 CACTGGAGGTCCTGCTGGGCTGTCCCGTTCCCCTGGGGGTGAAGCAGAAGCTTCTGCACT 2002
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4775 CACTGGAGGTCCTGCTGGGCTGTCCCGTTCCCCTGGGGGTGAAGCAGAAGCTTCTGCACT 4834

Qy      2003 GGGTCTCTCTGTTGGGTCAGCAGCCTAATGCCACCACCCCAGGAGACACCCTGGACGCCT 2062
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4835 GGGTCTCTCTGTTGGGTCAGCAGCCTAATGCCACCACCCCAGGAGACACCCTGGACGCCT 4894

Qy      2063 TCCACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAAACAGCTTCC 2122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4895 TCCACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTTCGCTTGGCATTAAACAGCTTCC 4954

Qy      2123 AAGAAGTGTGGCTTCCGATTAACCAGAACCTGGACTTGATAGCATCTTCCTTCTGCCTCC 2182
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4955 AAGAAGTGTGGCTTCCGATTAACCAGAACCTGGACTTGATAGCATCTTCCTTCTGCCTCC 5014

Qy      2183 AGCACTGTCCGTATTTGCGGAAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATG 2242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5015 AGCACTGTCCGTATTTGCGGAAAATTCGGGTGGATGTCAAAGGGATCTTCCCAAGAGATG 5074

Qy      2243 AGTCCGCTGAGGCATGTCCTGTGGTCCCTCTATGGATGCGGGATAAGACCCTCATTGAGG 2302
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2006, 23:24:38 ; Search time 645 Seconds
(without alignments)
10748.047 Million cell updates/sec

Title: US-10-677-943A-23
Perfect score: 3900
Sequence: 1 atgaaggttgcaggaggact.....gaaaaaaaaaaaaaaaaaa 3900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

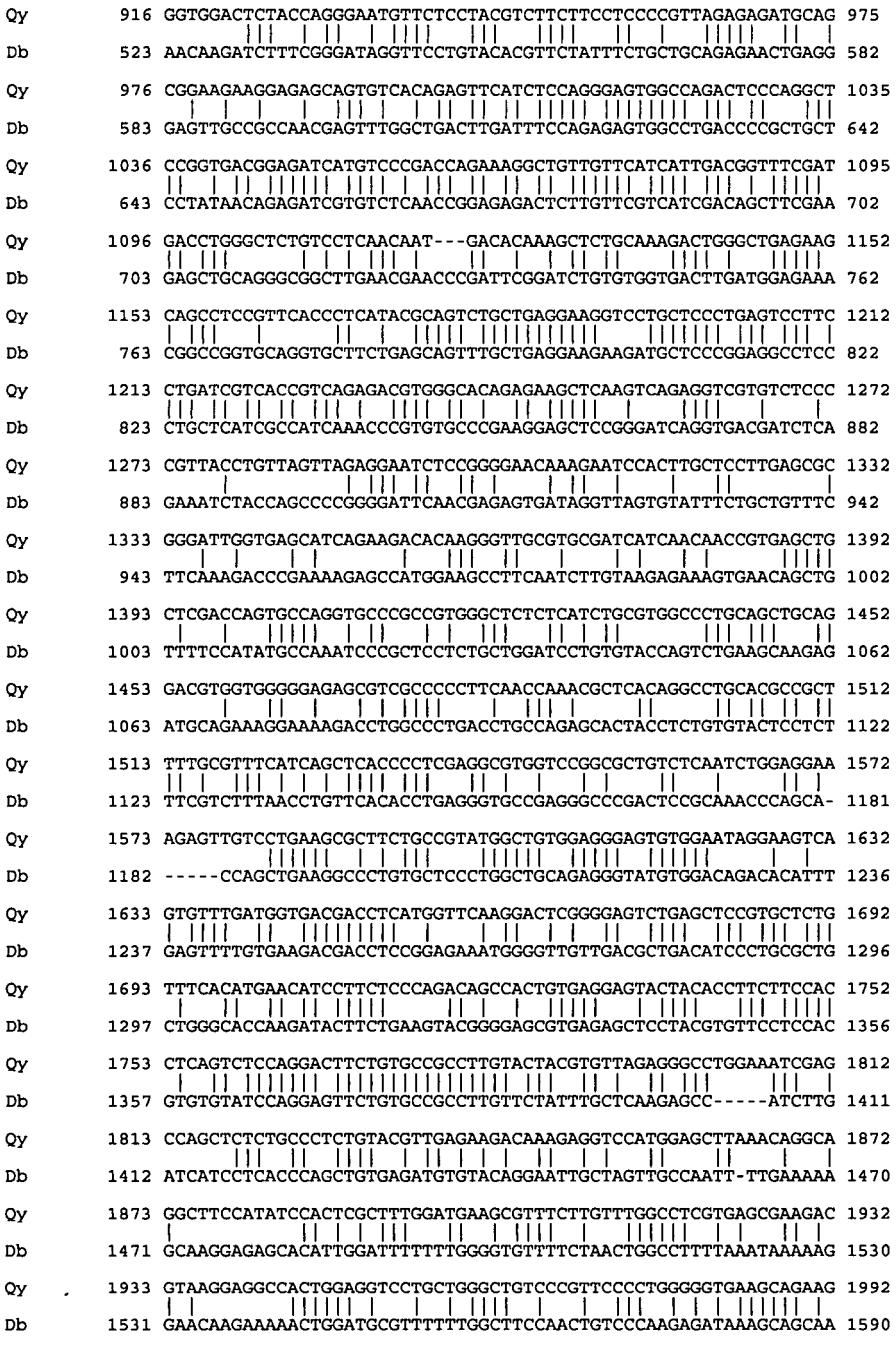
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	424	10.9	2985	3	US-10-781-294-15	Sequence 15, Appl
	2	401.2	10.3	3108	3	US-10-781-294-23	Sequence 23, Appl
	3	250.4	6.4	2524	3	US-10-781-294-58	Sequence 58, Appl
	4	219.8	5.6	4287	3	US-09-996-617-5	Sequence 5, Appli
	5	219.8	5.6	4422	3	US-09-388-221B-1	Sequence 1, Appli
	6	219.8	5.6	4556	3	US-09-388-221B-9	Sequence 9, Appli
	7	219.8	5.6	5444	3	US-09-996-617-1	Sequence 1, Appli
	8	218.4	5.6	4200	3	US-09-388-221B-3	Sequence 3, Appli
	9	218.4	5.6	4332	3	US-09-388-221B-5	Sequence 5, Appli
	10	218.4	5.6	4466	3	US-09-388-221B-11	Sequence 11, Appl
	11	182.4	4.7	1386	2	US-08-910-731-5	Sequence 5, Appli
	12	180.6	4.6	1371	2	US-08-910-731-1	Sequence 1, Appli
	13	180.6	4.6	1371	2	US-08-795-395-1	Sequence 1, Appli
	14	179.4	4.6	2046	3	US-10-781-294-19	Sequence 19, Appl
	15	179.2	4.6	1371	2	US-08-910-731-7	Sequence 7, Appli
	16	177.4	4.5	1374	2	US-08-910-731-3	Sequence 3, Appli
	17	177.4	4.5	1374	2	US-08-795-395-3	Sequence 3, Appli
	18	172.8	4.4	2844	3	US-10-781-294-17	Sequence 17, Appl
	19	151	3.9	2068	3	US-10-104-047-63	Sequence 63, Appl
	20	128.6	3.3	4931	3	US-10-183-770A-1	Sequence 1, Appli
	21	126.2	3.2	2689	3	US-10-029-347-1	Sequence 1, Appli
	22	126.2	3.2	2763	3	US-10-029-347-5	Sequence 5, Appli
	23	100.8	2.6	1515	3	US-10-781-294-21	Sequence 21, Appl
c	24	76.8	2.0	7218	2	US-08-232-463-14	Sequence 14, Appl
	25	60	1.5	376	3	US-09-023-655-395	Sequence 395, App
c	26	57	1.5	64309	3	US-09-949-016-14581	Sequence 14581, A
	27	54.2	1.4	2054	3	US-10-029-347-6	Sequence 6, Appli
	28	54.2	1.4	2223	2	US-08-257-073-4	Sequence 4, Appli
c	29	51.2	1.3	1037	3	US-09-181-585-3	Sequence 3, Appli
c	30	51.2	1.3	1159	3	US-09-181-585-1	Sequence 1, Appli
c	31	51.2	1.3	1471	3	US-09-181-585-2	Sequence 2, Appli
c	32	51.2	1.3	150394	3	US-09-949-016-13042	Sequence 13042, A
	33	50.2	1.3	7044	3	US-09-949-016-14113	Sequence 14113, A
	34	49.6	1.3	1276	3	US-09-177-325-2	Sequence 2, Appli
	35	49.6	1.3	1276	3	US-09-411-812A-2	Sequence 2, Appli
	36	49.6	1.3	1276	3	US-09-590-113-2	Sequence 2, Appli
	37	49	1.3	3123	3	US-10-002-974-53	Sequence 53, Appl
	38	49	1.3	3123	3	US-10-002-974-56	Sequence 56, Appl
	39	49	1.3	3123	3	US-10-002-974-58	Sequence 58, Appl
	40	49	1.3	3123	3	US-10-002-974-60	Sequence 60, Appl
	41	49	1.3	3123	3	US-10-002-974-62	Sequence 62, Appl
	42	49	1.3	3123	3	US-10-002-974-64	Sequence 64, Appl
	43	49	1.3	3123	3	US-10-002-974-66	Sequence 66, Appl
	44	49	1.3	3123	3	US-10-002-974-68	Sequence 68, Appl
	45	49	1.3	3123	3	US-10-002-974-84	Sequence 84, Appl

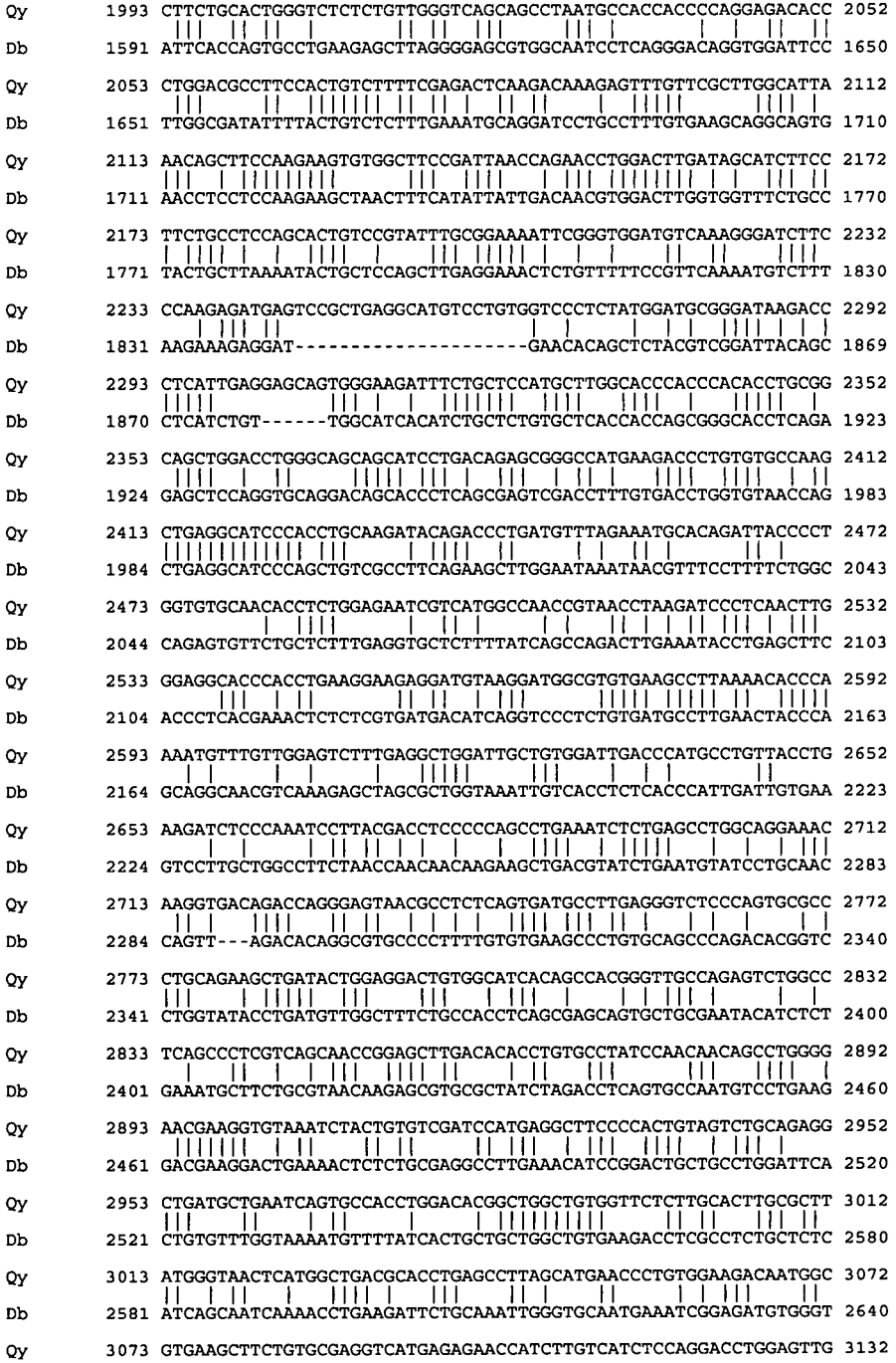
ALIGNMENTS

RESULT 1
US-10-781-294-15
; Sequence 15, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2985
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2985)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1857)
; OTHER INFORMATION: n = A,T,C or G
US-10-781-294-15

Query Match 10.9%; Score 424; DB 3; Length 2985;
Best Local Similarity 50.2%; Pred. No. 3.1e-113;
Matches 1262; Conservative 0; Mismatches 1205; Indels 45; Gaps 7;

```
Qy        796 TTGGCTGGTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGCACGGTGGTTCTGCAC 855
              ||||   |  ||||| | |  |       ||    | ||| || || ||| || | ||
Db        403 TTGGACCGCCTTTTTGCTCCCAAGGAAACTGGGAAACAGCCACGTACAGTGATTATTCAA 462

Qy        856 GGAAAGTCAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTGCTGTGCTGGGCGCAA 915
              |||     ||||||||| ||| || | ||    |  | | |  || ||  |||| || |
Db        463 GGACCACAAGGAATTGGAAAAACGACACTCCTGATGAAGCTGATGATGGCCTGGTCGGAC 522
```





GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2006, 23:33:05 ; Search time 2881 Seconds
(without alignments)
11194.229 Million cell updates/sec

Title: US-10-677-943A-23
Perfect score: 3900
Sequence: 1 atgaaggttgcaggaggact.....gaaaaaaaaaaaaaaaaaaa 3900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3900	100.0	3900	7	US-10-399-443-23	Sequence 23, Appl
2	3900	100.0	3900	7	US-10-677-943-23	Sequence 23, Appl
3	3885	99.6	3885	8	US-10-860-761-3	Sequence 3, Appli
4	3593.2	92.1	3926	6	US-10-216-645-1	Sequence 1, Appli
5	3559.2	91.3	3830	6	US-10-216-645-3	Sequence 3, Appli
6	3311.2	84.9	3489	7	US-10-416-642-3	Sequence 3, Appli
7	3166.2	81.2	4035	5	US-10-124-498-5	Sequence 5, Appli
8	3166.2	81.2	4035	5	US-10-066-521-5	Sequence 5, Appli
9	2965.2	76.0	3226	7	US-10-092-900A-347	Sequence 347, App
10	1292.8	33.1	3447	7	US-10-399-443-5	Sequence 5, Appli
11	1292.8	33.1	3447	7	US-10-677-943-5	Sequence 5, Appli
12	1149.6	29.5	1157	7	US-10-399-443-1	Sequence 1, Appli
13	1149.6	29.5	1157	7	US-10-677-943-1	Sequence 1, Appli
14	1073.4	27.5	1075	7	US-10-399-443-3	Sequence 3, Appli
15	1073.4	27.5	1075	7	US-10-677-943-3	Sequence 3, Appli
16	679.2	17.4	3521	8	US-10-794-342-2	Sequence 2, Appli
17	424	10.9	2985	3	US-09-965-621-15	Sequence 15, Appl
18	424	10.9	2985	6	US-10-407-866-15	Sequence 15, Appl
19	424	10.9	2985	7	US-10-781-294-15	Sequence 15, Appl
20	422.4	10.8	2982	5	US-10-124-498-25	Sequence 25, Appl
21	422.4	10.8	2982	5	US-10-066-521-25	Sequence 25, Appl
22	422.4	10.8	3190	6	US-10-094-749-1079	Sequence 1079, Ap
23	422.4	10.8	3368	5	US-10-124-498-23	Sequence 23, Appl
24	422.4	10.8	3368	5	US-10-066-521-23	Sequence 23, Appl
25	419.2	10.7	2575	3	US-09-848-035-7	Sequence 7, Appli
26	419.2	10.7	2575	3	US-09-986-224-7	Sequence 7, Appli
27	419.2	10.7	2575	9	US-10-116-422-7	Sequence 7, Appli
28	419.2	10.7	2883	7	US-10-161-493-3	Sequence 3, Appli
29	419.2	10.7	3172	7	US-10-398-037-74	Sequence 74, Appl
30	405.8	10.4	3186	5	US-10-124-498-17	Sequence 17, Appl
31	405.8	10.4	3186	5	US-10-066-521-17	Sequence 17, Appl
32	401.2	10.3	3108	3	US-09-965-621-23	Sequence 23, Appl
33	401.2	10.3	3108	6	US-10-407-866-23	Sequence 23, Appl
34	401.2	10.3	3108	7	US-10-781-294-23	Sequence 23, Appl
35	401.2	10.3	3218	6	US-10-407-866-67	Sequence 67, Appl
36	387.8	9.9	3531	9	US-10-756-149-3252	Sequence 3252, Ap
37	387.4	9.9	3540	8	US-10-794-342-7	Sequence 7, Appli
38	386.2	9.9	3189	6	US-10-239-663-10	Sequence 10, Appl
39	386.2	9.9	4944	9	US-10-499-353A-429	Sequence 429, App
40	382.8	9.8	2775	7	US-10-357-820-51	Sequence 51, Appl
c 41	380.6	9.8	1582	4	US-09-925-065A-74102	Sequence 74102, A
42	375	9.6	3099	5	US-10-127-516-3	Sequence 3, Appli
43	375	9.6	3099	5	US-10-027-629-3	Sequence 3, Appli
44	375	9.6	3099	6	US-10-132-967-3	Sequence 3, Appli
45	375	9.6	3431	5	US-10-127-516-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-399-443-23
; Sequence 23, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 3900
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3600)
; OTHER INFORMATION:
US-10-399-443-23

Query Match 100.0%; Score 3900; DB 7; Length 3900;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3900; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAAGGTTGCAGGAGGACTTGAACTTGGAGCTGCTGCTCTGCTCTCAGCATCACCACGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGGTTGCAGGAGGACTTGAACTTGGAGCTGCTGCTCTGCTCTCAGCATCACCACGT 60

Qy         61 GCTCTTGTCACTCTTTCCACAGGTCCTACTTGCTCTATATTACCAAAGAATCCACTTTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCTTGTCACTCTTTCCACAGGTCCTACTTGCTCTATATTACCAAAGAATCCACTTTTC 120

Qy        121 CCCCAAAACCTGAGCTCTCAGCCTTGTATCAAGATGGAAGGAGACAAATCGCTCACCTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      121 CCCCAAAACCTGAGCTCTCAGCCTTGTATCAAGATGGAAGGAGACAAATCGCTCACCTTT 180
Qy      181 TCCAGCTACGGGCTGCAATGGTGTCTCTATGAGCTAGACAAGGAAGAATTTCAGACATTC 240
Db      181 TCCAGCTACGGGCTGCAATGGTGTCTCTATGAGCTAGACAAGGAAGAATTTCAGACATTC 240
Qy      241 AAGGAATTACTAAAGAAGAAATCTTCAGAATCGACCACATGCTCTATTCCACAGTTTGAA 300
Db      241 AAGGAATTACTAAAGAAGAAATCTTCAGAATCGACCACATGCTCTATTCCACAGTTTGAA 300
Qy      301 ATCGAGAATGCCAACGTGGAATGTCTGGCACTCCTCTTGCATGAGTATTATGGAGCATCG 360
Db      301 ATCGAGAATGCCAACGTGGAATGTCTGGCACTCCTCTTGCATGAGTATTATGGAGCATCG 360
Qy      361 CTGGCCTGGGCTACGTCCATTAGCATCTTTGAAAACATGAACCTGCGAACCCTCTCGGAG 420
Db      361 CTGGCCTGGGCTACGTCCATTAGCATCTTTGAAAACATGAACCTGCGAACCCTCTCGGAG 420
Qy      421 AAGGCACGGGATGACATGAAAAGACATTCACCAGAAGATCCTGAAGCAACGATGACTGAC 480
Db      421 AAGGCACGGGATGACATGAAAAGACATTCACCAGAAGATCCTGAAGCAACGATGACTGAC 480
Qy      481 CAAGGACCAAGCAAGGAAAAAGTGCCAGGAATTTCACAAGCTGTGCAACAAGATAGTGCC 540
Db      481 CAAGGACCAAGCAAGGAAAAAGTGCCAGGAATTTCACAAGCTGTGCAACAAGATAGTGCC 540
Qy      541 ACAGCTGCAGAGACAAAAGAACAGGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACA 600
Db      541 ACAGCTGCAGAGACAAAAGAACAGGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACA 600
Qy      601 GCAGCAGAGACAGAAGAACAAGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCA 660
Db      601 GCAGCAGAGACAGAAGAACAAGAAATTTCACAAGCTATGGAACAAGAAGGTGCCACAGCA 660
Qy      661 GCAGAGACAGAAGAACAAGGACATGGAGGTGACACATGGGACTACAAGAGTCACGTGATG 720
Db      661 GCAGAGACAGAAGAACAAGGACATGGAGGTGACACATGGGACTACAAGAGTCACGTGATG 720
Qy      721 ACCAAATTCGCTGAGGAGGAGGATGTACGTCGTAGTTTTGAAAACACTGCTGCTGACTGG 780
Db      721 ACCAAATTCGCTGAGGAGGAGGATGTACGTCGTAGTTTTGAAAACACTGCTGCTGACTGG 780
Qy      781 CCGGAAATGCAAACGTTGGCTGGTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGC 840
Db      781 CCGGAAATGCAAACGTTGGCTGGTGCTTTTGATTCAGACCGGTGGGGCTTCCGGCCTCGC 840
Qy      841 ACGGTGGTTCTGCACGGAAAGTCAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTG 900
Db      841 ACGGTGGTTCTGCACGGAAAGTCAGGAATTGGGAAATCGGCTCTAGCCAGAAGGATCGTG 900
Qy      901 CTGTGCTGGGCGCAAGGTGGACTCTACCAGGGAATGTTCTCCTACGTCTTCTTCCTCCCC 960
Db      901 CTGTGCTGGGCGCAAGGTGGACTCTACCAGGGAATGTTCTCCTACGTCTTCTTCCTCCCC 960
Qy      961 GTTAGAGAGATGCAGCGGAAGAAGGAGAGCAGTGTCACAGAGTTCATCTCCAGGGAGTGG 1020
Db      961 GTTAGAGAGATGCAGCGGAAGAAGGAGAGCAGTGTCACAGAGTTCATCTCCAGGGAGTGG 1020
Qy     1021 CCAGACTCCCAGGCTCCGGTGACGGAGATCATGTCCCGACCAGAAAGGCTGTTGTTCATC 1080
Db     1021 CCAGACTCCCAGGCTCCGGTGACGGAGATCATGTCCCGACCAGAAAGGCTGTTGTTCATC 1080
Qy     1081 ATTGACGGTTTCGATGACCTGGGCTCTGTCCTCAACAATGACACAAAGCTCTGCAAAGAC 1140
Db     1081 ATTGACGGTTTCGATGACCTGGGCTCTGTCCTCAACAATGACACAAAGCTCTGCAAAGAC 1140
Qy     1141 TGGGCTGAGAAGCAGCCTCCGTTCACCCTCATACGCAGTCTGCTGAGGAAGGTCCTGCTC 1200
Db     1141 TGGGCTGAGAAGCAGCCTCCGTTCACCCTCATACGCAGTCTGCTGAGGAAGGTCCTGCTC 1200
Qy     1201 CCTGAGTCCTTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAG 1260
Db     1201 CCTGAGTCCTTCCTGATCGTCACCGTCAGAGACGTGGGCACAGAGAAGCTCAAGTCAGAG 1260
Qy     1261 GTCGTGTCTCCCCGTTACCTGTTAGTTAGAGGAATCTCCGGGGAACAAAGAATCCACTTG 1320
Db     1261 GTCGTGTCTCCCCGTTACCTGTTAGTTAGAGGAATCTCCGGGGAACAAAGAATCCACTTG 1320
Qy     1321 CTCCTTGAGCGCGGGATTGGTGAGCATCAGAAGACACAAGGGTTGCGTGCGATCATCAAC 1380
Db     1321 CTCCTTGAGCGCGGGATTGGTGAGCATCAGAAGACACAAGGGTTGCGTGCGATCATCAAC 1380
Qy     1381 AACCGTGAGCTGCTCGACCAGTGCCAGGTGCCCGCCGTGGGCTCTCTCATCTGCGTGGCC 1440
Db     1381 AACCGTGAGCTGCTCGACCAGTGCCAGGTGCCCGCCGTGGGCTCTCTCATCTGCGTGGCC 1440
Qy     1441 CTGCAGCTGCAGGACGTGGTGGGGGAGAGCGTCGCCCCCTTCAACCAAACGCTCACAGGC 1500
Db     1441 CTGCAGCTGCAGGACGTGGTGGGGGAGAGCGTCGCCCCCTTCAACCAAACGCTCACAGGC 1500
Qy     1501 CTGCACGCCGCTTTTGCGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTC 1560
Db     1501 CTGCACGCCGCTTTTGCGTTTCATCAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTC 1560
Qy     1561 AATCTGGAGGAAAGAGTTGTCCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGG 1620
Db     1561 AATCTGGAGGAAAGAGTTGTCCTGAAGCGCTTCTGCCGTATGGCTGTGGAGGGAGTGTGG 1620
Qy     1621 AATAGGAAGTCAGTGTTTGATGGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAG 1680
Db     1621 AATAGGAAGTCAGTGTTTGATGGTGACGACCTCATGGTTCAAGGACTCGGGGAGTCTGAG 1680
Qy     1681 CTCCGTGCTCTGTTTCACATGAACATCCTTCTCCCAGACAGCCACTGTGAGGAGTACTAC 1740
Db     1681 CTCCGTGCTCTGTTTCACATGAACATCCTTCTCCCAGACAGCCACTGTGAGGAGTACTAC 1740
Qy     1741 ACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCGCCTTGTACTACGTGTTAGAGGGC 1800
Db     1741 ACCTTCTTCCACCTCAGTCTCCAGGACTTCTGTGCCGCCTTGTACTACGTGTTAGAGGGC 1800
Qy     1801 CTGGAAATCGAGCCAGCTCTCTGCCCTCTGTACGTTGAGAAGACAAAGAGGTCCATGGAG 1860
Db     1801 CTGGAAATCGAGCCAGCTCTCTGCCCTCTGTACGTTGAGAAGACAAAGAGGTCCATGGAG 1860
Qy     1861 CTTAAACAGGCAGGCTTCCATATCCACTCGCTTTGGATGAAGCGTTTCTTGTTTGGCCTC 1920
Db     1861 CTTAAACAGGCAGGCTTCCATATCCACTCGCTTTGGATGAAGCGTTTCTTGTTTGGCCTC 1920
Qy     1921 GTGAGCGAAGACGTAAGGAGGCCACTGGAGGTCCTGCTGGGCTGTCCCGTTCCCCTGGGG 1980
Db     1921 GTGAGCGAAGACGTAAGGAGGCCACTGGAGGTCCTGCTGGGCTGTCCCGTTCCCCTGGGG 1980
Qy     1981 GTGAAGCAGAAGCTTCTGCACTGGGTCTCTCTGTTGGGTCAGCAGCCTAATGCCACCACC 2040
Db     1981 GTGAAGCAGAAGCTTCTGCACTGGGTCTCTCTGTTGGGTCAGCAGCCTAATGCCACCACC 2040
Qy     2041 CCAGGAGACACCCTGGACGCCTTCCACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTT 2100
Db     2041 CCAGGAGACACCCTGGACGCCTTCCACTGTCTTTTCGAGACTCAAGACAAAGAGTTTGTT 2100
Qy     2101 CGCTTGGCATTAAACAGCTTCCAAGAAGTGTGGCTTCCGATTAACCAGAACCTGGACTTG 2160
Db     2101 CGCTTGGCATTAAACAGCTTCCAAGAAGTGTGGCTTCCGATTAACCAGAACCTGGACTTG 2160
Qy     2161 ATAGCATCTTCCTTCTGCCTCCAGCACTGTCCGTATTTGCGGAAAATTCGGGTGGATGTC 2220
Db     2161 ATAGCATCTTCCTTCTGCCTCCAGCACTGTCCGTATTTGCGGAAAATTCGGGTGGATGTC 2220
Qy     2221 AAAGGGATCTTCCCAAGAGATGAGTCCGCTGAGGCATGTCCTGTGGTCCCTCTATGGATG 2280
Db     2221 AAAGGGATCTTCCCAAGAGATGAGTCCGCTGAGGCATGTCCTGTGGTCCCTCTATGGATG 2280
Qy     2281 CGGGATAAGACCCTCATTGAGGAGCAGTGGGAAGATTTCTGCTCCATGCTTGGCACCCAC 2340
Db     2281 CGGGATAAGACCCTCATTGAGGAGCAGTGGGAAGATTTCTGCTCCATGCTTGGCACCCAC 2340
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2006, 23:58:56 ; Search time 1975 Seconds
(without alignments)
7899.832 Million cell updates/sec

Title: US-10-677-943A-23
Perfect score: 3900
Sequence: 1 atgaaggttgcaggaggact.....gaaaaaaaaaaaaaaaaaaa 3900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq3:*
11: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
12: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
13: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

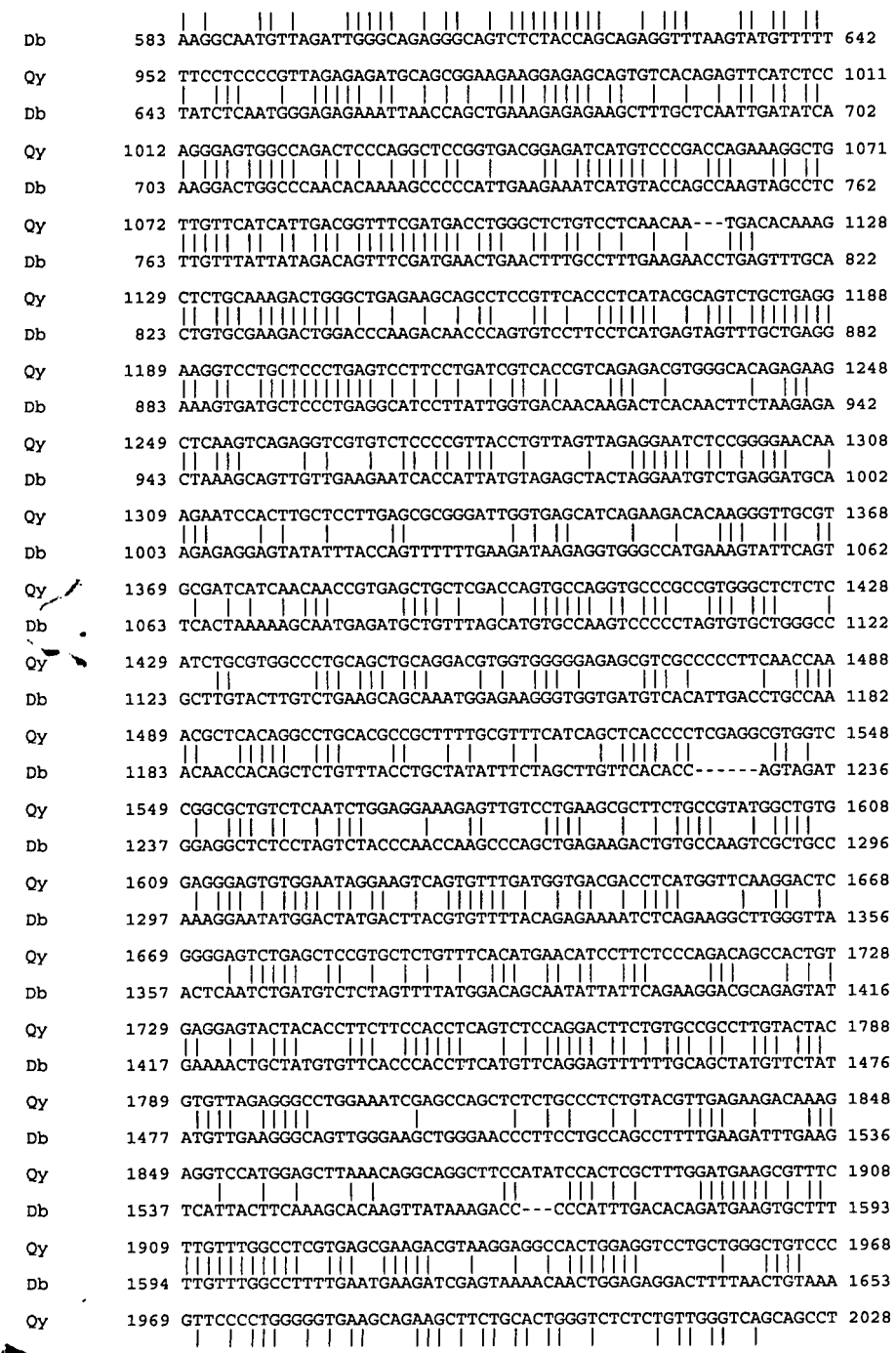
Result No.	Score	% Query Match	Length	DB	ID	Description
1	480.8	12.3	2808	9	US-10-511-989-13	Sequence 13, Appl
2	475.8	12.2	3039	9	US-10-511-989-11	Sequence 11, Appl
3	401.2	10.3	3731	9	US-10-511-989-1	Sequence 1, Appli
4	386.2	9.9	3189	14	US-11-137-465-10	Sequence 10, Appl
c 5	380.6	9.8	1582	6	US-09-925-065A-74102	Sequence 74102, A
c 6	380.6	9.8	1582	9	US-10-301-480-175341	Sequence 175341,
c 7	380.6	9.8	1582	10	US-10-301-480-788750	Sequence 788750,
8	363.4	9.3	3563	9	US-10-511-989-3	Sequence 3, Appli
9	346	8.9	3150	14	US-11-137-465-9	Sequence 9, Appli
10	326	8.4	2835	9	US-10-511-989-33	Sequence 33, Appl
11	320.4	8.2	3395	9	US-10-511-989-5	Sequence 5, Appli
12	319.4	8.2	3102	9	US-10-511-989-9	Sequence 9, Appli
13	304	7.8	2934	9	US-10-511-989-148	Sequence 148, App
14	263.6	6.8	3221	9	US-10-511-989-7	Sequence 7, Appli
15	248.8	6.4	2154	9	US-10-511-989-31	Sequence 31, Appl
16	221.4	5.7	5100	14	US-11-136-527-4000	Sequence 4000, Ap
17	212.4	5.4	1800	11	US-11-051-724-21	Sequence 21, Appl
18	192.2	4.9	1683	11	US-11-051-724-23	Sequence 23, Appl
19	179	4.6	993	9	US-10-511-989-35	Sequence 35, Appl
c 20	174	4.5	562	6	US-09-925-065A-817968	Sequence 817968,
21	174	4.5	2099	9	US-10-301-480-92963	Sequence 92963, A
22	174	4.5	2099	10	US-10-301-480-706372	Sequence 706372,
23	151	3.9	2068	11	US-11-072-512-63	Sequence 63, Appl
24	127.2	3.3	659	6	US-09-925-065A-752888	Sequence 752888,
25	101.8	2.6	536	14	US-11-128-061-276	Sequence 276, App
26	101.8	2.6	536	14	US-11-128-061-3918	Sequence 3918, Ap
27	101.8	2.6	536	14	US-11-128-049-276	Sequence 276, App
28	101.8	2.6	536	14	US-11-128-049-3918	Sequence 3918, Ap
c 29	94.8	2.4	1925	8	US-10-750-185-55918	Sequence 55918, A
c 30	94.8	2.4	1925	8	US-10-750-623-55918	Sequence 55918, A
c 31	93.8	2.4	1362	6	US-09-925-065A-713542	Sequence 713542,
32	88.2	2.3	1425	6	US-09-925-065A-57399	Sequence 57399, A
33	88.2	2.3	1425	6	US-09-925-065A-57400	Sequence 57400, A
34	88.2	2.3	1425	9	US-10-301-480-158637	Sequence 158637,
35	88.2	2.3	1425	9	US-10-301-480-158638	Sequence 158638,
36	88.2	2.3	1425	10	US-10-301-480-772046	Sequence 772046,
37	88.2	2.3	1425	10	US-10-301-480-772047	Sequence 772047,
38	80.4	2.1	925	10	US-10-301-480-611268	Sequence 611268,
39	80.4	2.1	925	10	US-10-301-480-1224677	Sequence 1224677,
40	78.4	2.0	613	6	US-09-925-065A-273421	Sequence 273421,
41	78.4	2.0	613	6	US-09-925-065A-273422	Sequence 273422,
42	78.4	2.0	622	10	US-10-301-480-350646	Sequence 350646,
43	78.4	2.0	622	10	US-10-301-480-350647	Sequence 350647,
44	78.4	2.0	622	10	US-10-301-480-964055	Sequence 964055,
45	78.4	2.0	622	10	US-10-301-480-964056	Sequence 964056,

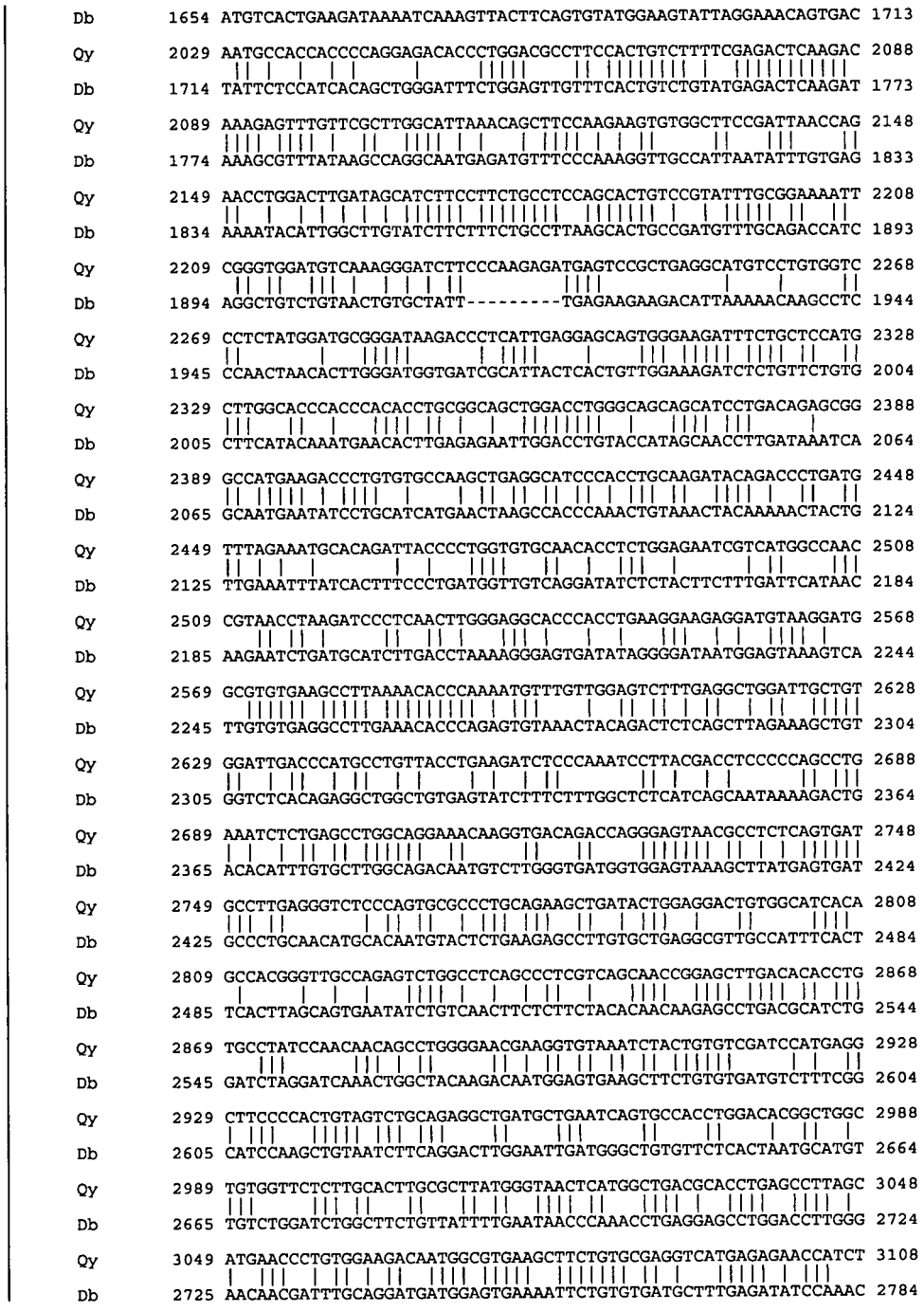
ALIGNMENTS

RESULT 1
US-10-511-989-13
; Sequence 13, Application US/10511989
; Publication No. US20060053496A1
; GENERAL INFORMATION:
; APPLICANT: University of North Carolina-Chapel Hill
; APPLICANT: Ting, Jenny
; APPLICANT: Linhoff, Michael
; APPLICANT: Harton, Johnathan
; APPLICANT: Williams, Kristi
; APPLICANT: Lich, John
; APPLICANT: O'Connor, William
; APPLICANT: Moore, Christopher
; APPLICANT: Davis, Beckley
; APPLICANT: Brickey, W. Jane
; APPLICANT: Conti, Brian
; APPLICANT: Zhang, Jinghua
; APPLICANT: Zhu, Xin-Sheng
; TITLE OF INVENTION: CATERPILLER GENE FAMILY
; FILE REFERENCE: 5470.368WO
; CURRENT APPLICATION NUMBER: US/10/511,989
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 60/376,626
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 2808
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-989-13

Query Match 12.3%; Score 480.8; DB 9; Length 2808;
Best Local Similarity 51.9%; Pred. No. 2.3e-117;
Matches 1194; Conservative 0; Mismatches 1087; Indels 21; Gaps 4;

```
Qy      832 CGGCCTCGCACGGTGGTTCTGCACGGAAAGTCAGGAATTGGGAAATCGGCTCTAGCCAGA 891
            | ||| |  |  ||||| || || |||    | ||| |||||||| |  |  | |  |||
Db      523 CAGCCACAGATCGTGGTGCTTCAGGGAGCTGCTGGAGTTGGGAAAACAACCTTGGTGAGA 582

Qy      892 AGGATCGTGCTGTGCTGGGCGCAAGGTGGACTCTACCAGGGAATGTTCTCCTACGTCTTC 951
```





GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2006, 12:13:56 ; Search time 194 Seconds
(without alignments)
2717.807 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6298	100.0	1200	5	AAU79526	Aau79526 Human MAT
2	6298	100.0	1200	6	AAE31749	Aae31749 Human MAT
3	6298	100.0	1200	9	ADV04681	Adv04681 NALP5, SE
4	6077.5	96.5	2312	5	AAO17859	Aao17859 Pyrin dom
5	5892.5	93.6	1162	5	AAE25053	Aae25053 Human EMB
6	5892.5	93.6	1162	6	ADA45219	Ada45219 Human MAT
7	5804	92.2	1143	6	ADA45221	Ada45221 Human MAT
8	5741.5	91.2	1344	5	AAO15585	Aao15585 Human PYR
9	5741.5	91.2	1344	7	ADF94745	Adf94745 Human PYR
10	5134	81.5	1033	5	ABU65214	Abu65214 Human NOV
11	5125	81.4	1033	8	ADN62079	Adn62079 Human nov
12	3553	56.4	682	8	ADJ19405	Adj19405 Human PAN
13	2747	43.6	1111	5	AAU79525	Aau79525 Mouse MAT
14	2610	41.4	1111	6	AAE31748	Aae31748 Mouse MAT
15	2484	39.4	476	8	ADP47894	Adp47894 Human CAT
16	1997	31.7	385	5	AAU79523	Aau79523 Human MAT
17	1932	30.7	385	6	AAE31746	Aae31746 Human MAT
18	1874.5	29.8	1093	9	ADV44750	Adv44750 Human nuc
19	1659.5	26.3	1012	8	ADP47734	Adp47734 Human inf
20	1516.5	24.1	935	8	ADP47736	Adp47736 Human inf
21	1409.5	22.4	994	5	AAO15593	Aao15593 Human PYR
22	1409.5	22.4	994	5	ADE36409	Ade36409 Human PAA
23	1409.5	22.4	994	7	ADF94761	Adf94761 Human PYR
24	1409.5	22.4	994	8	ADJ19331	Adj19331 Human PAN
25	1408.5	22.4	994	8	ABM82183	Abm82183 Tumour-as
26	1397	22.2	924	7	ADF44705	Adf44705 Human NOV
27	1387	22.0	924	8	ADO50890	Ado50890 Human NOV
28	1379	21.9	258	5	AAU79524	Aau79524 Human MAT
29	1369	21.7	983	5	AAO17870	Aao17870 Pyrin dom
30	1322.5	21.0	1062	8	ADP47724	Adp47724 Human Mon
31	1320	21.0	991	9	ADV44751	Adv44751 Human nuc
32	1320	21.0	1035	5	ADE36417	Ade36417 Human PAA
33	1320	21.0	1035	8	ADJ19339	Adj19339 Human PAN
34	1320	21.0	1099	5	AAO17857	Aao17857 Pyrin dom
35	1319	20.9	1027	8	ADJ19383	Adj19383 Human PAN
36	1319	20.9	1061	5	AAO15590	Aao15590 Human PYR
37	1319	20.9	1061	7	ADF94755	Adf94755 Human PYR
38	1312	20.8	1036	9	ADV44755	Adv44755 Human nuc
39	1304.5	20.7	919	6	ADA55150	Ada55150 Human pro
40	1303.5	20.7	919	5	ABP43483	Abp43483 Human sec
41	1303.5	20.7	919	6	ABJ19319	Abj19319 NOVX rela
42	1303.5	20.7	919	8	ADO41638	Ado41638 Novel hum
43	1301	20.7	258	6	AAE31747	Aae31747 Human MAT
44	1278.5	20.3	858	5	AAM50328	Aam50328 Human nuc
45	1252.5	19.9	1006	8	ADP47726	Adp47726 Human Mon

ALIGNMENTS

RESULT 1
AAU79526
ID AAU79526 standard; protein; 1200 AA.
XX
AC AAU79526;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human MATER protein.
XX
KW Human; contraceptive; antiinfertility; MATER;
KW maternal antigen that embryos require; MATER null phenotype; oocyte;
KW early embryonic survival; premature ovarian failure; POF;
KW autoimmune infertility; chromosome 19; gene therapy; fertility.
XX
OS Homo sapiens.
XX
PN WO200232955-A1.
XX
PD 25-APR-2002.
XX
PF 04-APR-2001; 2001WO-US010981.
XX
PR 18-OCT-2000; 2000US-0241510P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson LM, Tong Z;
XX
DR WPI; 2002-454595/48.
DR N-PSDB; ABK48628.
XX
PT New isolated human Maternal Antigen That Embryos Require protein and
PT polynucleotide, useful in diagnosing or treating fertility or reduced
PT fertility, or as a contraceptive.
XX
PS Claim 1; Fig 9; 93pp; English.
XX
CC The invention discloses an isolated human MATER (Maternal Antigen That
CC Embryos Require) protein, which can complement a MATER null phenotype in
CC which zygotes arising from the oocyte do not progress beyond the two-cell
CC stage. MATER is required for early embryonic survival and abnormal levels
CC of the protein can lead to premature ovarian failure (POF) and can be
CC caused by under or over expression of MATER or an autoimmune response to
CC MATER. MATER is a single-copy maternal effect gene found on chromosome
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in
CC diagnosing or treating fertility and reduced fertility. In particular,

```
CC   the MATER protein is useful as a contraceptive agent, or for influencing
CC   (either inhibiting or enhancing) fertility and can be used to detect a
CC   predisposition to infertility or reduced fertility, or for presymptomatic
CC   screening of an individual for infertility/reduced fertility. The protein
CC   and polynucleotide are also useful for detecting an excess or deficiency,
CC   or genetic mutation, of the MATER protein in a mammalian subject (e.g. a
CC   human or a mouse) or for screening for a compound useful in influencing
CC   MATER-mediated fertility. The protein sequence presented is the human
CC   MATER. Note: Residues 1-25 are encoded by nucleotides which were deduced,
CC   by the inventors, by comparison to published human genomic DNA sequences,
CC   while the remainder of the sequence was determined by direct cloning of
CC   human ovarian cDNAs
XX
SQ   Sequence 1200 AA;

  Query Match           100.0%;  Score 6298;  DB 5;  Length 1200;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60

Qy         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120

Qy        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180

Qy        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240

Qy        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300

Qy        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360

Qy        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420

Qy        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480

Qy        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540

Qy        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600

Qy        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660

Qy        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720

Qy        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780

Qy        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840

Qy        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900

Qy        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960

Qy        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020

Qy       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080

Qy       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140

Qy       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200


RESULT 2
AAE31749
ID   AAE31749 standard; protein; 1200 AA.
XX
AC   AAE31749;
XX
DT   07-MAR-2003  (first entry)
XX
DE   Human MATER protein.
XX
KW   Human; MATER protein; infertility; fertility; contraceptive agent;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200281492-A1.
XX
PD   17-OCT-2002.
XX
PF   29-MAR-2002; 2002WO-US009776.
XX
PR   04-APR-2001; 2001WO-US010981.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Nelson LM,  Tong Z;
XX
DR   WPI; 2003-058494/05.
DR   N-PSDB; AAD49018.
XX
PT   New isolated variant MATER proteins and nucleic acids, useful for
PT   diagnosing, prognosing and treating infertility and reduced fertility,
PT   and as contraceptive agents.
XX
PS   Claim 1; Page 99-104; 110pp; English.
XX
CC   The present invention relates to novel MATER proteins and polynucleotides
CC   encoding such proteins. The MATER proteins are essential to fertility.
CC   Sequences of the invention are useful for diagnosing, prognosing and
CC   treating infertility, reduced fertility and as contraceptive agents. They
CC   are also useful in gene therapy. The method is useful for detecting a
CC   predisposition to or pre-symptomatic screening of an individual for
CC   infertility or reduced fertility. The present sequence is human MATER
CC   protein
XX
SQ   Sequence 1200 AA;
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2006, 12:17:51 ; Search time 50 Seconds
(without alignments)
2309.202 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2747	43.6	1111	2	A59000	mater protein [imp
2	914	14.5	1192	2	T17255	hypothetical prote
3	633.5	10.1	456	2	S20597	ribonuclease inhib
4	624	9.9	456	2	A31857	ribonuclease inhib
5	611.5	9.7	461	2	A31858	ribonuclease-angio
6	367.5	5.8	483	2	S27880	Nasopressin recept
7	327	5.2	1130	2	A48843	MHC class II trans
8	220	3.5	631	2	C89243	protein F28C1.3 [i
9	220	3.5	631	2	T21471	hypothetical prote
10	205	3.3	1075	2	T31668	hypothetical prote
11	201.5	3.2	545	2	T52068	RAN GTPase-activat
12	197	3.1	1004	2	T31665	hypothetical prote
13	192.5	3.1	568	2	F86291	hypothetical prote
14	188.5	3.0	1167	2	T06146	disease resistance
15	185	2.9	533	2	T52063	ran GTPase-activat
16	183	2.9	535	2	T48102	RAN GTPase activat
17	179.5	2.9	607	2	E96598	protein F20N2.2 [i
18	177	2.8	789	2	T52067	hypothetical prote
19	175.5	2.8	506	2	A45841	T-complex-associat
20	172	2.7	907	2	F96617	probable disease r
21	172	2.7	1039	2	A85096	hypothetical prote
22	171	2.7	934	2	T05201	hypothetical prote
23	168.5	2.7	1184	2	H71436	hypothetical prote
24	168.5	2.7	1301	2	D85188	disease resistance
25	167	2.7	712	2	C71419	hypothetical prote
26	164.5	2.6	1063	2	E96662	hypothetical prote
27	163.5	2.6	1317	2	B85189	disease resistance
28	162.5	2.6	998	2	T23427	hypothetical prote
29	161	2.6	1422	2	B71437	probable resistanc
30	159.5	2.5	1253	2	T45787	disease resistance
31	158.5	2.5	1240	2	T06404	resistance complex
32	157.5	2.5	1152	2	T31911	hypothetical prote
33	155.5	2.5	1219	2	T06608	disease resistance
34	153.5	2.4	618	2	T48193	hypothetical prote
35	151.5	2.4	1239	2	T06143	disease resistence
36	150.5	2.4	630	2	S13724	ESAG 8 protein - T
37	150.5	2.4	1017	2	H96663	hypothetical prote
38	150	2.4	271	2	D84586	hypothetical prote
39	148.5	2.4	587	2	JC5300	Ran GTPase activat
40	148	2.3	1304	2	G85188	disease resistance
41	147.5	2.3	966	2	D96662	hypothetical prote
42	147	2.3	589	2	T52070	RNA1 protein homol
43	146	2.3	589	2	A36983	RNA1 homolog fug1
44	145	2.3	1121	2	T02764	myosin-I binding p
45	145	2.3	1189	2	T52346	disease resistance

ALIGNMENTS

RESULT 1
A59000
mater protein [imported] - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: A59000
R;Tong, Z.B.; Nelson, L.M.
Endocrinology 140, 3720-3726, 1999
A;Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature ov
A;Reference number: A59000; MUID:99360614; PMID:10433232
A;Accession: A59000
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1111 <KUR>
A;Cross-references: UNIPROT:Q9R1M5; UNIPARC:UPI0000021410; GB:AF074018; NID:g5802697; PI
C;Genetics:
A;Gene: Mater

Query Match 43.6%; Score 2747; DB 2; Length 1111;
Best Local Similarity 50.7%; Pred. No. 2.6e-176;
Matches 564; Conservative 185; Mismatches 307; Indels 56; Gaps 14;

```
Qy      133 MNLRTLSEKARDDMKRHSPE-DPEATMTDQGPSKE-----------------KVPGISQ 173
            :  | | |: : : |  ||| | ::   |||| :|                 |  |: :
Db       12 LKARGLEEEQKSERKMTSPENDSKSIQKDQGPEQEQTSESTMGPPEKDSKAILKARGLEE 71

Qy      174 AVQQDS----------ATAAETKEQEISQAMEQEGATAAETEEQEISQ----AMEQEGAT 219
              : :|           |   ::  :|: || |::  |: | : : | :      ||   |
Db       72 EQKSESTMSPSENVSRAILKDSGSEEVEQASERK-MTSPENDSKSIQKDQGPEQEQTSET 130

Qy      220 AAETEE-------QGHGGDTWDYKSHVMTKFAEEEDVRRSFENTAADWPEMQTLAGAFDS 272
            ||       : :|||  |||:||: ||    |:         | |||: |: ||
Db      131 LQSKEEDEVTEADKDNGGDLQDYKAHVIAKFDTSVDLH-------YDSPEMKLLSDAFKP 183

Qy      273 DRWGFRPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSV 332
             :  |:| |::|||: |:|||||| ||| |||| |:| | |:| |  |||::  ::||:
Db      184 YQKTFQPHTIILHGRPGVGKSALARSIVLGWAQGKLFQKM-SFVIFFSVREIKWTEKSSL 242

Qy      333 TEFISREWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVL-NNDTKLCKDWAEKQPPFTLI 391
             : |::| |||   ||:|||:|||||:||| ||: ||| ::|  | :|| ::|| : |:
Db      243 AQLIAKECPDSWDLVTKIMSQPERLLFVIDGLDDMDSVLQHDDMTLSRDWKDEQPIYILM 302

Qy      392 RSLLRKVLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQK 451
            ||||| |||:||||:| |: | ||||| |||| |:|| |:|  :|  |:||     |  :
Db      303 YSLLRKALLPQSFLIITTRNTGLEKLKSMVVSPLYILVEGLSASRRSQLVLENISNESDR 362

Qy      452 TQGLRAIINNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLT 511
            |   ::| | :| |||| |:| ||:| ||||  :|:      ||||||:|   |||||
Db      363 IQVFHSLIENHQLFDQCQAPSVCSLVCEALQLQKKLGKRCTLPCQTLTGLYATLVFHQLT 422

Qy      512 PRGVVRRCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILL 571
             :   :  |: ||:: |   | || ||||  :|||  |||    | |||: |||||||||
```

```
Db        423 LKRPSQSALSQEEQITLVGLCMMAAEGVWTMRSVFYDDDLKNYSLKESEILALFHMNILL 482

Qy        572 PDSH-CEEYYTFFHLSLQDFCAALYYVLEGL-EIEPALCPLYVEKTKRSMELKQAGFHIH 629

Db        483 QVGHNSEQCYVFSHLSLQDFFAALYYVLEGLEEWNQHFC--FIENQRSIMEVKRTD-DTR 539

Qy        630 SLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFH 689

Db        540 LLGMKRFLFGLMNKDILKTLEVLFEYPVIPTVEQKLQHWVSLIAQQVNGTSPMDTLDAFY 599

Qy        690 CLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDES 749

Db        600 CLFESQDEEFVGGALKRFQEVWLLINQKMDLKVSSYCLKHCQNLKAIRVDIRDLLSVDNT 659

Qy        750 AEACPVVPLW-MRDKTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHP 808

Db        660 LELCPVVTVQETQCKPLLMEWWGNFCSVLGSLRNLKELDLGDSILSQRAMKILCLELRNQ 719

Qy        809 TCKIQTLMFRNAQITPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLL 868

Db        720 SCRIQKLTFKSAEVVSGLKHLWKLLFSNQNLKYLNLGNTPMKDDDMKLACEALKHPKCSV 779

Qy        869 ESLRLDCCGLTHACYLKISQILTTSPSLKSLSLAGNKVTDQGVTPLSDALRVSQCALQKL 928

Db        780 ETLRLDSCELTIIGYEMISTLLISTTRLKCLSLAKNRVGVKSMISLGNALSSSMCLLQKL 839

Qy        929 ILEDCGITATGCQSLASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLN 988

Db        840 ILDNCGLTPASCHLLVSALFSNQNLTHLCLSNNSLGTEGVQQLCQFLRNPECALQRLILN 899

Qy        989 QCHLDTAGCGSLALALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCH 1048

Db        900 HCNIVDDAYGFLAMRLANNTKLTHLSLTMNPVGDGAMKLLCEALKEPTCYLQELELVDCQ 959

Qy       1049 LTAACCESLSCVISRSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTS 1108

Db        960 LTQNCCEDLACMITTTKHLKSLDLGNNALGDKGVITLCEGLKQSSSSLRRLGLGACKLTS 1019

Qy       1109 DCCEALSLALSCNRHLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQYPVQIR 1168

Db       1020 NCCEALSLAISCNPHLNSLNLVKNDFSTSGMLKLCSAFQCPVSNLGIIGLWKQEYYARVR 1079

Qy       1169 KLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200

Db       1080 RQLEEVEFVKPHVVIDGDWYASDEDDRNWWKN 1111
```

RESULT 2
T17255
hypothetical protein DKFZp586O1822.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17255
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18722
A;Accession: T17255
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1192 <KOE>
A;Cross-references: UNIPROT:Q9C000; UNIPARC:UPI000016AC73; EMBL:AL117470
A;Experimental source: adult uterus; clone DKFZp586O1822
C;Genetics:
A;Note: DKFZp586O1822.1

  Query Match 14.5%; Score 914; DB 2; Length 1192;
  Best Local Similarity 27.3%; Pred. No. 5.1e-53;
  Matches 274; Conservative 142; Mismatches 314; Indels 274; Gaps 28;

```
Qy        249 VRRSFENTAADWP-----------EMQTLAG-AFDSDRWGFRPRTVVLHGKSGIGKSALA 296

Db         15 VKRS-------WPDYVEENRGHLIEIRDLFGPGLDTQ----EPRIVILQGAAGIGKSTLA 63
```

```
Qy        297 RRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPER 356

Db         64 RQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPER 123

Qy        357 LLFIIDGFDDLGSVLNN-DTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTE 415

Db        124 LLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQ 183

Qy        416 KLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGS 475

Db        184 NLIPSLEQARWVEVLGFSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSW 243

Qy        476 LICVALQLQDVVGESVAPFNQTLTG--LHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCR 533

Db        244 LACTCLMQQMKRKEKLTLTSKTTTTLCLHYLAQALQAQPLG-----------PQLRDLCS 292

Qy        534 MAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNIL----LPDSHCEEYYTFFHLSLQD 589

Db        293 LAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHPIPLS-----YSFIHLCFQE 347

Qy        590 FCAALYYVLEGLE--IEPALCPLYVEKTKRSMELKQAGFHIHSLW---MKRFLFGLVSED 644

Db        348 FFAAMSYVLEDEKGRGKHSNCIIDLEKTLEA-------YGIHGLFGASTTRFLLGLLSDE 400

Qy        645 VRRPLEVLLGCPVPLGVKQKLLHWVSLLG--QQPNATTPGDTLDAFHCLFETQDKEFVRL 702

Db        401 GEREMENIFHCRLSQG--RNLMQWVPSLQLLLQPH------SLESLHCLYETRNKTFLTQ 452

Qy        703 ALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRD 762

Db        453 VMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQL-IEG---RQHRSTWSPTMVVLFRW 508

Qy        763 KTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQI 822

Db        509 VPVTDAYW---------------------------------------------------- 516

Qy        823 TPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHAC 882

Db        517 ----QILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLT--- 569

Qy        883 YLKISQILTTSPSLKSLSLAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQS 942

Db        570 ----------------------------------------------AED---------- 572

Qy        943 LASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 1002

Db        573 ----------------------------------------------------CKDLAF 578

Qy       1003 ALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVIS 1062

Db        579 GLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLS 638

Qy       1063 RSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSL------ 1116

Db        639 ASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRALEQEKP 698

Qy       1117 -ALSCNRHLTSL---------NLVQNNFSPKGMMKLCSAFAC---PTSNLQIIGLWKWQY 1163

Db        699 QLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERAASHVAQANLKLLDVSK-IF 757

Qy       1164 PVQIRKLLEE-------VQLL-KPRVVIDGSWHS--FDEDDRYW 1197

Db        758 P--IAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 799
```

RESULT 3
S20597
ribonuclease inhibitor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S20597
R;Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
Biochim. Biophys. Acta 1129, 335-338, 1992

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2006, 12:14:26 ; Search time 258 Seconds
(without alignments)
3281.527 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6298	100.0	1200	1	NALP5_HUMAN	P59047 homo sapien
2	3106.5	49.3	1098	2	Q647I9_BOVIN	Q647i9 bos taurus
3	2754	43.7	1111	2	Q6XZF2_MOUSE	Q6xzf2 mus musculu
4	2753	43.7	1095	2	Q6VSG3_MOUSE	Q6vsg3 mus musculu
5	2747	43.6	1111	1	NALP5_MOUSE	Q9r1m5 mus musculu
6	2747	43.6	1111	2	Q53ZD5_MOUSE	Q53zd5 mus musculu
7	2747	43.6	1111	2	Q7TPU9_MOUSE	Q7tpu9 mus musculu
8	2746	43.6	1095	2	Q6VSG5_MOUSE	Q6vsg5 mus musculu
9	2745	43.6	1111	2	Q6VSG0_MOUSE	Q6vsg0 mus musculu
10	2626	41.7	1085	2	Q6VSG1_MOUSE	Q6vsg1 mus musculu
11	2262	35.9	989	2	Q6VSG4_MOUSE	Q6vsg4 mus musculu
12	1874.5	29.8	1093	1	NAL14_HUMAN	Q86w24 homo sapien
13	1525	24.2	727	2	Q6VSG2_MOUSE	Q6vsg2 mus musculu
14	1515	24.1	727	2	Q6VSF9_MOUSE	Q6vsf9 mus musculu
15	1417.5	22.5	993	2	Q6B966_MOUSE	Q6b966 mus musculu
16	1409.5	22.4	994	1	NALP4_HUMAN	Q96mn2 homo sapien
17	1408.5	22.4	994	2	Q86W87_HUMAN	Q86w87 homo sapien
18	1404.5	22.3	986	2	Q66X14_MOUSE	Q66x14 mus musculu
19	1355	21.5	996	2	Q4U4S5_BOVIN	Q4u4s5 bos taurus
20	1330	21.1	1043	1	NAL13_HUMAN	Q86w25 homo sapien
21	1320	21.0	991	2	Q7RTR0_HUMAN	Q7rtr0 homo sapien
22	1319	20.9	1061	1	NAL12_HUMAN	P59046 homo sapien
23	1289.5	20.5	986	2	Q86W27_HUMAN	Q86w27 homo sapien
24	1252.5	19.9	1034	1	CIAS1_HUMAN	Q96p20 homo sapien
25	1250	19.8	1033	1	CIAS1_MOUSE	Q8r4b8 mus musculu
26	1245.5	19.8	982	2	Q66X27_MOUSE	Q66x27 mus musculu
27	1242	19.7	1062	2	Q53FL5_HUMAN	Q53fl5 homo sapien
28	1241	19.7	1062	1	NALP2_HUMAN	Q9nx02 homo sapien
29	1226.5	19.5	978	2	Q66X19_MOUSE	Q66x19 mus musculu
30	1178	18.7	982	2	Q8BU40_MOUSE	Q8bu40 mus musculu
31	1145	18.2	1048	2	Q7RTR4_HUMAN	Q7rtr4 homo sapien
32	1128	17.9	1003	2	Q66X22_MOUSE	Q66x22 mus musculu
33	1115	17.7	1009	2	Q7RTR1_HUMAN	Q7rtr1 homo sapien
34	1097	17.4	980	1	NALP7_HUMAN	Q8wx94 homo sapien
35	1082.5	17.2	949	2	Q66X03_MOUSE	Q66x03 mus musculu
36	1082.5	17.2	1029	2	Q86W28_HUMAN	Q86w28 homo sapien
37	1060	16.8	1033	1	NAL11_HUMAN	P59045 homo sapien
38	1060	16.8	1033	2	Q53ZZ0_HUMAN	Q53zz0 homo sapien
39	1047	16.6	1004	2	Q66X01_MOUSE	Q66x01 mus musculu
40	983	15.6	1473	1	NALP1_HUMAN	Q9c000 homo sapien
41	981	15.6	1375	2	Q86UB5_HUMAN	Q86ub5 homo sapien
42	954.5	15.2	959	2	Q66X05_MOUSE	Q66x05 mus musculu
43	951	15.1	934	2	Q66X24_MOUSE	Q66x24 mus musculu
44	944.5	15.0	931	2	Q6UQE7_MOUSE	Q6uqe7 mus musculu
45	940	14.9	863	2	Q8C6J9_MOUSE	Q8c6j9 mus musculu

ALIGNMENTS

RESULT 1
NALP5_HUMAN
ID NALP5_HUMAN STANDARD; PRT; 1200 AA.
AC P59047; Q86W29;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 5 (Mater protein homolog).
GN Name=NALP5; Synonyms=MATER;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RX MEDLINE=21922687; PubMed=11925379; DOI=10.1093/humrep/17.4.903;
RA Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
RT "A human homologue of mouse Mater, a maternal effect gene essential
RT for early embryonic development.";
RL Hum. Reprod. 17:903-911(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22451042; PubMed=12563287; DOI=10.1038/nrm1019;
RA Tschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
CC -!- TISSUE SPECIFICITY: Oocyte specific.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 13 LRR (leucine-rich) repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- CAUTION: It is not obvious that this is the ortholog of mouse
CC Mater.
CC ---
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ---
DR EMBL; AY054986; AAL15549.1; -; mRNA.
DR EMBL; AY154460; AAO18156.1; -; mRNA.
DR HSSP; P13489; 1A4Y.
DR Ensembl; ENSG00000171487; Homo sapiens.
DR HGNC; HGNC:21269; NALP5.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN.
DR Pfam; PF00560; LRR_1; 2.
DR Pfam; PF05729; NACHT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS50824; DAPIN; 1.

```
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Nucleotide-binding; Repeat.
FT   DOMAIN       57    148       DAPIN.
FT   DOMAIN      280    602       NACHT.
FT   REPEAT      704    727       LRR 1.
FT   REPEAT      730    753       LRR 2.
FT   REPEAT      780    803       LRR 3.
FT   REPEAT      809    832       LRR 4.
FT   REPEAT      836    863       LRR 5.
FT   REPEAT      865    892       LRR 6.
FT   REPEAT      893    916       LRR 7.
FT   REPEAT      950    973       LRR 8.
FT   REPEAT      979   1002       LRR 9.
FT   REPEAT     1007   1034       LRR 10.
FT   REPEAT     1036   1059       LRR 11.
FT   REPEAT     1064   1092       LRR 12.
FT   REPEAT     1121   1142       LRR 13.
FT   NP_BIND     286    293       ATP (Potential).
FT   CONFLICT    459    459       I -> M (in Ref. 2).
FT   CONFLICT    506    506       A -> V (in Ref. 2).
FT   CONFLICT    912    912       T -> M (in Ref. 2).
FT   CONFLICT    999    999       S -> F (in Ref. 2).
SQ   SEQUENCE   1200 AA;  134236 MW;  9A070D2A771B28FA CRC64;

  Query Match             100.0%;  Score 6298;  DB 1;  Length 1200;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60

Qy         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120

Qy        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180

Qy        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240

Qy        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300

Qy        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360

Qy        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPPTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPPTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420

Qy        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480

Qy        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540

Qy        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600

Qy        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660

Qy        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720

Qy        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780

Qy        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840

Qy        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900

Qy        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960

Qy        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020

Qy       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080

Qy       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140

Qy       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200


RESULT 2
Q647I9_BOVIN
ID   Q647I9_BOVIN PRELIMINARY;      PRT;  1098 AA.
AC   Q647I9;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Maternal antigen that embryo require.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC   Pecora; Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15189828; DOI=10.1095/biolreprod.104.030288;
RA   Pennetier S., Uzbekova S., Perreau C., Papillier P., Mermillod P.,
RA   Dalbies-Tran R.;
RT   "Spatio-temporal expression of the germ cell marker genes MATER, ZAR1,
RT   GDF9, BMP15,andVASA in adult bovine tissues, oocytes, and
RT   preimplantation embryos.";
RL   Biol. Reprod. 71:1359-1366(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Pennetier S., Uzbekova S., Dalbies-Tran R.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY721594; AAU20764.1; -; mRNA.
DR   InterPro; IPR000887; Aldlse_KDPG_KHG.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN.
DR   Pfam; PF00560; LRR_1; 1.
DR   Pfam; PF05729; NACHT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00368; LRR_RI; 12.
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2006, 12:22:16 ; Search time 51 Seconds
(without alignments)
1945.310 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSPDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

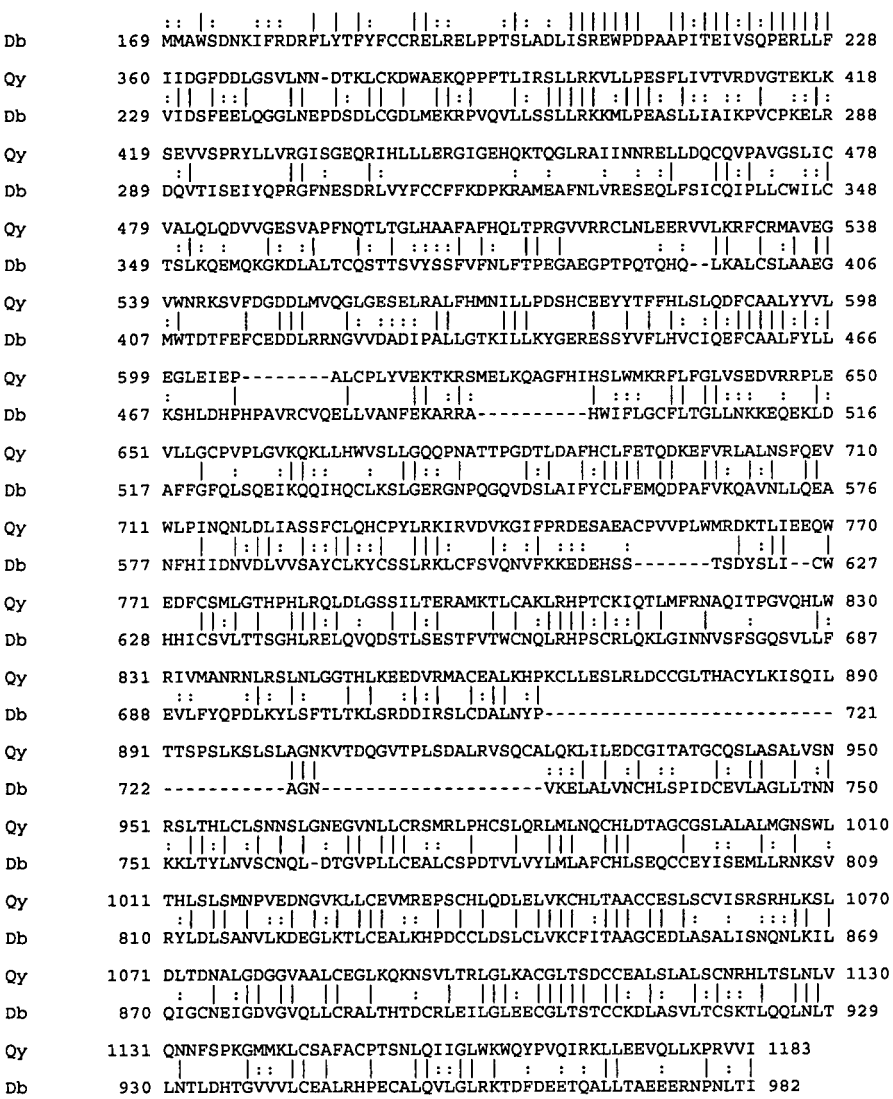
Result No.	Score	% Query Match	Length	DB	ID	Description
1	1409.5	22.4	994	2	US-10-781-294-16	Sequence 16, Appl
2	1320	21.0	1035	2	US-10-781-294-24	Sequence 24, Appl
3	1252.5	19.9	1034	2	US-10-183-770A-18	Sequence 18, Appl
4	1224.5	19.4	1033	2	US-10-029-347-4	Sequence 4, Appli
5	985	15.6	1429	2	US-10-029-347-3	Sequence 3, Appli
6	985	15.6	1429	2	US-10-183-770A-3	Sequence 3, Appli
7	983	15.6	1429	2	US-10-029-347-26	Sequence 26, Appl
8	983	15.6	1429	2	US-10-183-770A-15	Sequence 15, Appl
9	983	15.6	1429	2	US-09-996-617-2	Sequence 2, Appli
10	983	15.6	1473	2	US-09-388-221B-2	Sequence 2, Appli
11	979	15.5	1454	2	US-09-388-221B-10	Sequence 10, Appl
12	976	15.5	1399	2	US-09-388-221B-4	Sequence 4, Appli
13	976	15.5	1443	2	US-09-388-221B-6	Sequence 6, Appli
14	972	15.4	1424	2	US-09-388-221B-12	Sequence 12, Appl
15	876.5	13.9	603	2	US-10-781-294-64	Sequence 64, Appl
16	713.5	11.3	947	2	US-10-781-294-18	Sequence 18, Appl
17	693	11.0	719	2	US-10-781-294-59	Sequence 59, Appl
18	633.5	10.1	456	1	US-08-910-731-4	Sequence 4, Appli
19	633.5	10.1	456	1	US-08-795-395-4	Sequence 4, Appli
20	624	9.9	456	1	US-08-910-731-2	Sequence 2, Appli
21	624	9.9	456	1	US-08-795-395-2	Sequence 2, Appli
22	624	9.9	456	2	US-10-188-495-64	Sequence 64, Appl
23	615	9.8	456	1	US-08-910-731-8	Sequence 8, Appli
24	611.5	9.7	681	2	US-10-781-294-20	Sequence 20, Appl
25	610.5	9.7	461	1	US-08-910-731-6	Sequence 6, Appli
26	590	9.4	514	2	US-10-029-347-25	Sequence 25, Appl
27	590	9.4	514	2	US-10-183-770A-14	Sequence 14, Appl
28	582	9.2	375	2	US-10-781-294-39	Sequence 39, Appl
29	556.5	8.8	449	2	US-10-183-770A-2	Sequence 2, Appli
30	541.5	8.6	516	2	US-10-104-047-2033	Sequence 2033, Ap
31	529.5	8.4	319	2	US-10-781-294-37	Sequence 37, Appl
32	507.5	8.1	625	2	US-10-029-347-2	Sequence 2, Appli
33	390.5	6.2	317	2	US-10-781-294-60	Sequence 60, Appl
34	383.5	6.1	504	2	US-10-781-294-22	Sequence 22, Appl
35	373	5.9	270	2	US-10-781-294-38	Sequence 38, Appl
36	369.5	5.9	953	2	US-09-245-281-43	Sequence 43, Appl
37	369.5	5.9	953	2	US-09-207-359B-43	Sequence 43, Appl
38	369.5	5.9	953	2	US-09-340-620A-43	Sequence 43, Appl
39	369.5	5.9	953	2	US-09-865-364-43	Sequence 43, Appl
40	369.5	5.9	953	2	US-09-728-721-43	Sequence 43, Appl
41	356	5.7	1040	2	US-10-002-974-59	Sequence 59, Appl
42	356	5.7	1040	2	US-10-002-974-85	Sequence 85, Appl
43	354.5	5.6	209	2	US-09-388-221B-19	Sequence 19, Appl
44	354	5.6	1013	2	US-10-014-269-3	Sequence 3, Appli
45	354	5.6	1013	2	US-10-002-974-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-781-294-16
; Sequence 16, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 994
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-781-294-16

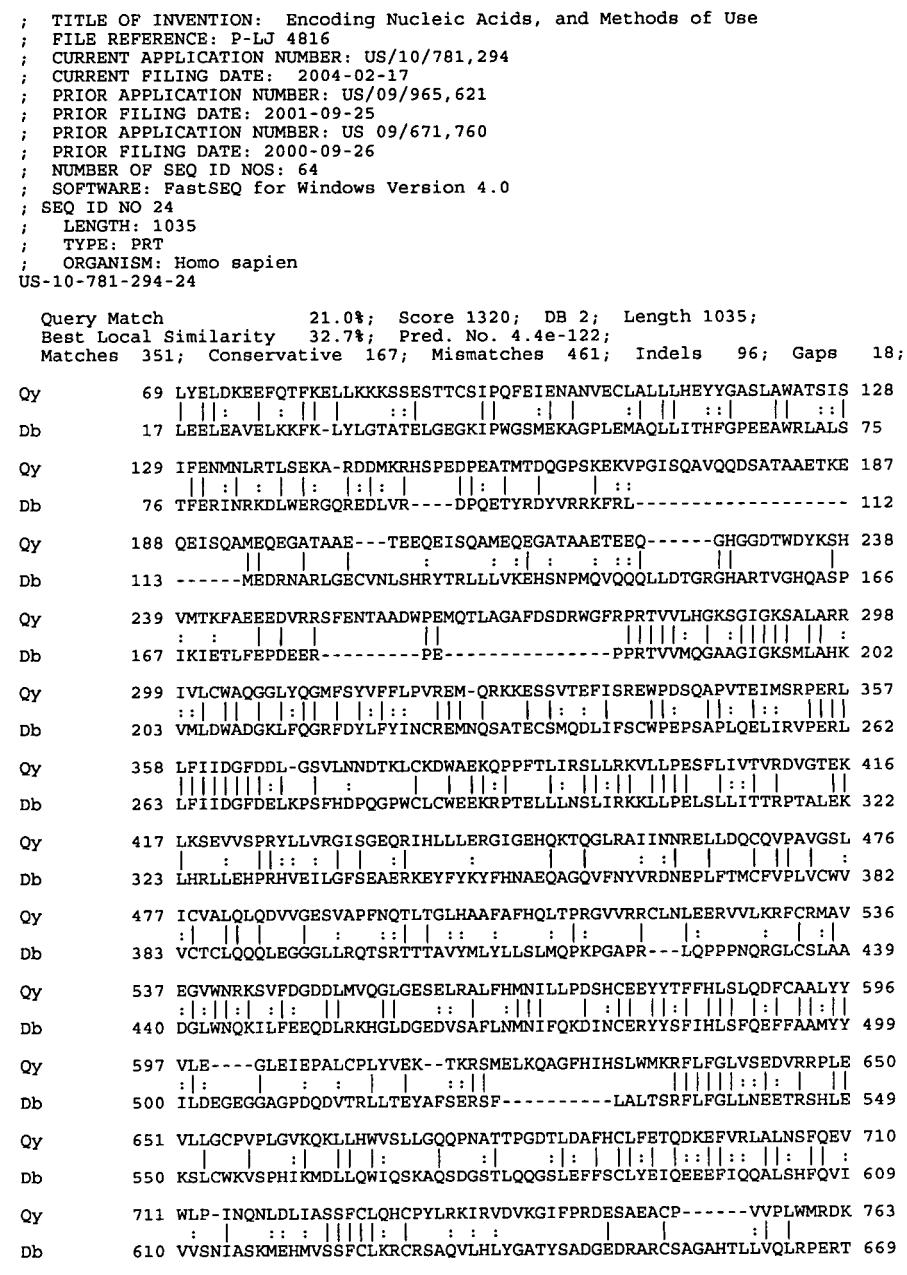
Query Match 22.4%; Score 1409.5; DB 2; Length 994;
Best Local Similarity 30.0%; Pred. No. 4.4e-131;
Matches 340; Conservative 199; Mismatches 429; Indels 165; Gaps 12;

```
Qy          60 FSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGA 119
               || :|| | | || ||||: ||| ||: : :     ||  |:: |: | || || ::|
Db           6 FSDFGLMWYLEELKKEEFRKFKEHLKQMTLQLELKQIPWTEVKKASREELANLLIKHYEE 65

Qy         120 SLAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDS 179
                 ||  :: ||: |: : |  |   :  :       | | |  :|:|   :
Db          66 QQAWNITLRIFQKMDRKDLCMKVMRERTGY-------TKTYQAHAKQKFSRL-------- 110

Qy         180 ATAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHV 239
                                                                | ||
Db         111 ----------------------------------------------------WSSKSVT 117

Qy         240 MTKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRI 299
                  ||:|::          |  |   |     | :||||:: |  |||||: |  ::
Db         118 EIHLYFEEEVKQE---------ECDHLDRLFAPKETGKQPRTVIIQGPQGIGKTTLLMKL 168

Qy         300 VLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLF 359
```



RESULT 2
US-10-781-294-24
; Sequence 24, Application US/10781294
; Patent No. 6953691
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-781-294-24

Query Match 21.0%; Score 1320; DB 2; Length 1035;
Best Local Similarity 32.7%; Pred. No. 4.4e-122;
Matches 351; Conservative 167; Mismatches 461; Indels 96; Gaps 18;



GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2006, 12:33:47 ; Search time 180 Seconds
(without alignments)
2785.529 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6298	100.0	1200	4	US-10-399-443-24	Sequence 24, Appl
2	6298	100.0	1200	4	US-10-677-943-24	Sequence 24, Appl
3	6298	100.0	1200	5	US-10-860-761-4	Sequence 4, Appli
4	5892.5	93.6	1162	4	US-10-216-645-2	Sequence 2, Appli
5	5892.5	93.6	1162	4	US-10-416-642-1	Sequence 1, Appli
6	5804	92.2	1143	4	US-10-216-645-4	Sequence 4, Appli
7	5741.5	91.2	1344	4	US-10-124-498-6	Sequence 6, Appli
8	5741.5	91.2	1344	4	US-10-066-521-6	Sequence 6, Appli
9	5134	81.5	1033	4	US-10-092-900A-348	Sequence 348, App
10	3553	56.4	682	4	US-10-407-866-96	Sequence 96, Appl
11	2747	43.6	1111	4	US-10-216-645-5	Sequence 5, Appli
12	2610	41.4	1111	4	US-10-399-443-6	Sequence 6, Appli
13	2610	41.4	1111	4	US-10-677-943-6	Sequence 6, Appli
14	1932	30.7	385	4	US-10-399-443-2	Sequence 2, Appli
15	1932	30.7	385	4	US-10-677-943-2	Sequence 2, Appli
16	1874.5	29.8	1093	5	US-10-794-342-13	Sequence 13, Appl
17	1409.5	22.4	994	3	US-09-965-621-16	Sequence 16, Appl
18	1409.5	22.4	994	4	US-10-124-498-24	Sequence 24, Appl
19	1409.5	22.4	994	4	US-10-066-521-24	Sequence 24, Appl
20	1409.5	22.4	994	4	US-10-407-866-16	Sequence 16, Appl
21	1409.5	22.4	994	4	US-10-781-294-16	Sequence 16, Appl
22	1397	22.2	924	4	US-10-357-820-52	Sequence 52, Appl
23	1320	21.0	991	5	US-10-794-342-14	Sequence 14, Appl
24	1320	21.0	1035	3	US-09-965-621-24	Sequence 24, Appl
25	1320	21.0	1035	4	US-10-407-866-24	Sequence 24, Appl
26	1320	21.0	1035	4	US-10-781-294-24	Sequence 24, Appl
27	1319	20.9	1027	4	US-10-407-866-68	Sequence 68, Appl
28	1319	20.9	1061	4	US-10-124-498-18	Sequence 18, Appl
29	1319	20.9	1061	4	US-10-066-521-18	Sequence 18, Appl
30	1312	20.8	1036	5	US-10-794-342-18	Sequence 18, Appl
31	1304.5	20.7	919	4	US-10-094-749-2718	Sequence 2718, Ap
32	1303.5	20.7	919	4	US-10-161-493-4	Sequence 4, Appli
33	1303.5	20.7	919	4	US-10-398-037-7	Sequence 7, Appli
34	1301	20.7	258	4	US-10-399-443-4	Sequence 4, Appli
35	1301	20.7	258	4	US-10-677-943-4	Sequence 4, Appli
36	1278.5	20.3	858	3	US-09-848-035-8	Sequence 8, Appli
37	1278.5	20.3	858	3	US-09-986-224-8	Sequence 8, Appli
38	1278.5	20.3	858	5	US-10-116-422-8	Sequence 8, Appli
39	1252.5	19.9	1032	5	US-10-882-761-33	Sequence 33, Appl
40	1252.5	19.9	1034	4	US-10-127-516-5	Sequence 5, Appli
41	1252.5	19.9	1034	4	US-10-027-629-5	Sequence 5, Appli
42	1252.5	19.9	1034	4	US-10-028-374-18	Sequence 18, Appl
43	1252.5	19.9	1034	4	US-10-132-967-5	Sequence 5, Appli
44	1252.5	19.9	1034	4	US-10-183-770-18	Sequence 18, Appl
45	1252.5	19.9	1034	4	US-10-264-958B-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-10-399-443-24
; Sequence 24, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-399-443-24

Query Match 100.0%; Score 6298; DB 4; Length 1200;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60

Qy         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120

Qy        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180

Qy        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240

Qy        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300

Qy        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360

Qy        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420

Qy        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480

Qy        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540

Qy        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600

Qy        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660

Qy        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720

Qy        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780

Qy        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLWRIVMANRNLR 840

Qy        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 SLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHACYLKISQILTTSPSLKSLS 900

Qy        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSNRSLTHLCLSN 960

Qy        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 NSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLALALMGNSWLTHLSLSMNPV 1020

Qy       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVISRSRHLKSLDLTDNALGDG 1080

Qy       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCNRHLTSLNLVQNNFSPKGMM 1140

Qy       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KLCSAFACPTSNLQIIGLWKWQYPVQIRKLLEEVQLLKPRVVIDGSWHSFDEDDRYWWKN 1200


RESULT 2
US-10-677-943-24
; Sequence 24, Application US/10677943
; Publication No. US20040072297A1
; GENERAL INFORMATION:
;  APPLICANT: The Government of the United States of America as
;  APPLICANT:  represented by the Secretary of the Department of Health and
;  APPLICANT:  Human Services
;  APPLICANT:  Nelson, Lawrence
;  APPLICANT:  Tong, Zhi-Bin
;  TITLE OF INVENTION: Human Gene Critical to Fertility
;  FILE REFERENCE: 4239-64790
;  CURRENT APPLICATION NUMBER: US/10/677,943
;  CURRENT FILING DATE:  2003-10-01
;  PRIOR APPLICATION NUMBER: 60/241,510
;  PRIOR FILING DATE: 2000-10-18
;  PRIOR APPLICATION NUMBER: PCT/US02/09776
;  PRIOR FILING DATE: 2002-03-29
;  PRIOR APPLICATION NUMBER: PCT/US01/10981
;  PRIOR FILING DATE: 2001-04-04
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
;   LENGTH: 1200
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-677-943-24

  Query Match           100.0%;  Score 6298;  DB 4;  Length 1200;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKVAGGLELGAAALLSASPRALVTLSTGPTCSILPKNPLFPQNLSSQPCIKMEGDKSLTF 60

Qy         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALLLHEYYGAS 120

Qy        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGISQAVQQDSA 180

Qy        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDTWDYKSHVM 240

Qy        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TKFAEEEDVRRSFENTAADWPEMQTLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIV 300

Qy        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFI 360

Qy        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IDGFDDLGSVLNNDTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSE 420

Qy        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVA 480

Qy        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVW 540

Qy        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEG 600

Qy        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 LEIEPALCPLYVEKTKRSMELKQAGFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLG 660

Qy        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDL 720

Qy        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 IASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQWEDFCSMLGTH 780
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2006, 12:35:07 ; Search time 26 Seconds
(without alignments)
1439.630 Million cell updates/sec

Title: US-10-677-943A-24
Perfect score: 6298
Sequence: 1 MKVAGGLELGAAALLSASPR.....VVIDGSWHSFDEDDRYWWKN 1200

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /SIDS5/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /SIDS5/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2484	39.4	476	6	US-10-511-989-173	Sequence 173, App
2	1659.5	26.3	1012	6	US-10-511-989-12	Sequence 12, Appl
3	1516.5	24.1	935	6	US-10-511-989-14	Sequence 14, Appl
4	1322.5	21.0	1062	6	US-10-511-989-2	Sequence 2, Appli
5	1252.5	19.9	1006	6	US-10-511-989-4	Sequence 4, Appli
6	1241	19.7	1062	7	US-11-137-465-43	Sequence 43, Appl
7	1197.5	19.0	1049	7	US-11-137-465-42	Sequence 42, Appl
8	1196.5	19.0	1033	6	US-10-511-989-10	Sequence 10, Appl
9	1190.5	18.9	944	6	US-10-511-989-34	Sequence 34, Appl
10	1160	18.4	977	6	US-10-511-989-149	Sequence 149, App
11	1138.5	18.1	950	6	US-10-511-989-6	Sequence 6, Appli
12	1048	16.6	892	6	US-10-511-989-8	Sequence 8, Appli
13	805.5	12.8	582	7	US-11-051-724-22	Sequence 22, Appl
14	756.5	12.0	474	6	US-10-511-989-167	Sequence 167, App
15	755	12.0	717	6	US-10-511-989-32	Sequence 32, Appl
16	745.5	11.8	521	7	US-11-051-724-24	Sequence 24, Appl
17	705.5	11.2	472	6	US-10-511-989-168	Sequence 168, App
18	702	11.1	472	6	US-10-511-989-163	Sequence 163, App
19	680	10.8	479	6	US-10-511-989-162	Sequence 162, App
20	677	10.7	477	6	US-10-511-989-169	Sequence 169, App
21	626.5	9.9	458	6	US-10-511-989-166	Sequence 166, App
22	618	9.8	466	6	US-10-511-989-172	Sequence 172, App
23	613	9.7	468	6	US-10-511-989-164	Sequence 164, App
24	587.5	9.3	1065	6	US-10-511-989-28	Sequence 28, Appl
25	556	8.8	496	6	US-10-511-989-174	Sequence 174, App
26	541.5	8.6	516	7	US-11-072-512-2033	Sequence 2033, Ap
27	535	8.5	473	6	US-10-511-989-170	Sequence 170, App
28	518.5	8.2	330	6	US-10-511-989-36	Sequence 36, Appl
29	516	8.2	467	6	US-10-511-989-175	Sequence 175, App
30	437.5	6.9	470	6	US-10-511-989-171	Sequence 171, App
31	398.5	6.3	1154	6	US-10-511-989-22	Sequence 22, Appl
32	391	6.2	932	6	US-10-511-989-20	Sequence 20, Appl
33	389	6.2	950	6	US-10-511-989-18	Sequence 18, Appl
34	387	6.1	1866	6	US-10-511-989-186	Sequence 186, App
35	380.5	6.0	454	6	US-10-511-989-176	Sequence 176, App
36	370.5	5.9	870	6	US-10-511-989-16	Sequence 16, Appl
37	350	5.6	953	6	US-10-966-846-2	Sequence 2, Appli
38	318	5.0	1122	6	US-10-330-773-260	Sequence 260, App
39	307.5	4.9	1487	6	US-10-511-989-24	Sequence 24, Appl
40	297.5	4.7	297	6	US-10-511-989-165	Sequence 165, App
41	269.5	4.3	588	6	US-10-511-989-177	Sequence 177, App
42	267.5	4.2	743	6	US-10-511-989-26	Sequence 26, Appl
43	266.5	4.2	501	6	US-10-511-989-180	Sequence 180, App
44	260.5	4.1	1057	6	US-10-330-773-253	Sequence 253, App
45	259	4.1	1070	7	US-11-147-047-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-10-511-989-173
; Sequence 173, Application US/10511989
; Publication No. US20060053496A1
; GENERAL INFORMATION:
; APPLICANT: University of North Carolina-Chapel Hill
; APPLICANT: Ting, Jenny
; APPLICANT: Linhoff, Michael
; APPLICANT: Harton, Johnathan
; APPLICANT: Williams, Kristi
; APPLICANT: Lich, John
; APPLICANT: O'Connor, William
; APPLICANT: Moore, Christopher
; APPLICANT: Davis, Beckley
; APPLICANT: Brickey, W. Jane
; APPLICANT: Conti, Brian
; APPLICANT: Zhang, Jinghua
; APPLICANT: Zhu, Xin-Sheng
; TITLE OF INVENTION: CATERPILLER GENE FAMILY
; FILE REFERENCE: 5470.368WO
; CURRENT APPLICATION NUMBER: US/10/511,989
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 60/376,626
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 173
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-989-173

Query Match 39.4%; Score 2484; DB 6; Length 476;
Best Local Similarity 99.6%; Pred. No. 1.2e-188;
Matches 474; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy        265 TLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREM 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TLAGAFDSDRWGFRPRTVVLHGKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREM 60

Qy        325 QRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEK 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QRKKESSVTEFISREWPDSQAPVTEIMSRPERLLFIIDGFDDLGSVLNNDTKLCKDWAEK 120

Qy        385 QPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLER 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLER 180
```

```
Qy        445 GIGEHQKTQGLRAIINNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAA 504
Db        181 GIGEHQKTQGLRAIMNNRELLDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAA 240

Qy        505 FAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRAL 564
Db        241 FVFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRAL 300

Qy        565 FHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQA 624
Db        301 FHMNILLPDSHCEEYYTFFHLSLQDFCAALYYVLEGLEIEPALCPLYVEKTKRSMELKQA 360

Qy        625 GFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDT 684
Db        361 GFHIHSLWMKRFLFGLVSEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDT 420

Qy        685 LDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDV 740
Db        421 LDAFHCLFETQDKEFVRLALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDV 476
```

RESULT 2
US-10-511-989-12
; Sequence 12, Application US/10511989
; Publication No. US20060053496A1
; GENERAL INFORMATION:
; APPLICANT: University of North Carolina-Chapel Hill
; APPLICANT: Ting, Jenny
; APPLICANT: Linhoff, Michael
; APPLICANT: Harton, Johnathan
; APPLICANT: Williams, Kristi
; APPLICANT: Lich, John
; APPLICANT: O'Connor, William
; APPLICANT: Moore, Christopher
; APPLICANT: Davis, Beckley
; APPLICANT: Brickey, W. Jane
; APPLICANT: Conti, Brian
; APPLICANT: Zhang, Jinghua
; APPLICANT: Zhu, Xin-Sheng
; TITLE OF INVENTION: CATERPILLER GENE FAMILY
; FILE REFERENCE: 5470.368WO
; CURRENT APPLICATION NUMBER: US/10/511,989
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 60/376,626
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-989-12

Query Match 26.3%; Score 1659.5; DB 6; Length 1012;
Best Local Similarity 34.9%; Pred. No. 7.5e-123;
Matches 391; Conservative 189; Mismatches 406; Indels 135; Gaps 16;

```
Qy         53 EGDKSLTFSSYGLQWCLYELDKEEFQTFKELLKKKSSESTTCSIPQFEIENANVECLALL 112
Db          3 DSSSSSFFPDFGLLLYLEELNKEELNTFK-LFLKETMEPEHGLTPWNEVKKARREDLANL 61

Qy        113 LHEYYGASLAWATSISIFENMNLRTLSEKARDDMKRHSPEDPEATMTDQGPSKEKVPGIS 172
Db         62 MKKYYPGEKAWSVSLKIFGKMNLKDLCERA----------------------KEEINWSA 99

Qy        173 QAVQQDSATAAETKEQEISQAMEQEGATAAETEEQEISQAMEQEGATAAETEEQGHGGDT 232
Db        100 QTIGPDDAKAGETQE-------------------------DQEAVL----------GDG 123

Qy        233 WDYKSHVMTKFAEEEDVRRSFENTAADW------PEMQTLAGAFDSD-RWGFRPRTVVLH 285
Db        124 TEYRNRIKEKFCITWD-KKSLAGKPEDFHHGIAEKDRKLLEHLFDVDVKTGAQPQIVVLQ 182

Qy        286 GKSGIGKSALARRIVLCWAQGGLYQGMFSYVFFLPVREMQRKKESSVTEFISREWPDSQA 345
Db        183 GAAGVGKTTLVRKAMLDWAEGSLYQQRFKYVFYLNGREINQLKERSFAQLISKDWPSTEG 242

Qy        346 PVTEIMSRPERLLFIIDGFDDLGSVLNN-DTKLCKDWAEKQPPFTLIRSLLRKVLLPESF 404
Db        243 PIEEIMYQPSSLLFIIDSFDELNFAFEEPEFALCEDWTQEHPVSFLMSSLLRKVMLPEAS 302

Qy        405 LIVTVRDVGTEKLKSEVVSPRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNREL 464
Db        303 LLVTTRLTTSKRLKQLLKNHHYVELLGMSEDAREEYIYQFFEDKRWAMKVFSSLKSNEML 362

Qy        465 LDQCQVPAVGSLICVALQLQDVVGESVAPFNQTLTGLHAAFAFHQLTPRGVVRRCLNLEE 524
Db        363 FSMCQVPLVCWAACTCLKQQMEKGGDVTLTCQTTTALFTCYISSLFTP--VDGGSPSLPN 420

Qy        525 RVVLKRFCRMAVEGVWNRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEEYYTFFH 584
Db        421 QAQLRRLCQVAAKGIWTMTYVFYRENLRRLGLTQSDVSSFMDSNIIQKDAEYENCYVFTH 480

Qy        585 LSLQDFCAALYYVLEG-LEIEPALCPLY--VEKTKRSMELKQAGFHIHSLWMKRFLFGLV 641
Db        481 LHVQEFFAAMFYMLKGSWEAGNPSCQPFEDLKSLLQSTSYKDP----HLTQMKCFLFGLL 536

Qy        642 SEDVRRPLEVLLGCPVPLGVKQKLLHWVSLLGQQPNATTPGDTLDAFHCLFETQDKEFVR 701
Db        537 NEDRVKQLERTFNCKMSLKIKSKLLQCMEVLGNSDYSPSQLGFLELFHCLYETQDKAFIS 596

Qy        702 LALNSFQEVWLPINQNLDLIASSFCLQHCPYLRKIRVDVKGIFPRDESAEACPVVPLWMR 761
Db        597 QAMRCFPKVAINICEKIHLLVSSFCLKHCRCLRTIRLSVTVVFEKKILKTSLP-TNTWLK 655

Qy        762 DKTLIEEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQ 821
Db        656 FIT-----FPDGCQ-----------DISTSLI---------------------------- 671

Qy        822 ITPGVQHLWRIVMANRNLRSLNLGGTHLKEEDVRMACEALKHPKCLLESLRLDCCGLTHA 881
Db        672 ------------HNKNLMHLDLKGSDIGDNGVKSLCEALKHPECKLQTLRLESCNLTVF 718

Qy        882 CYLKISQILTTSPSLKSLSLAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQ 941
Db        719 CCLNISNALIRSQSLIFLNLSTNNLLDDGVQLLCEALRHPKCYLERLSLESCGLTEAGCE 778

Qy        942 SLASALVSNRSLTHLCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLA 1001
Db        779 YLSLALISNKRLTHLCLADNVLGDGGVKLMSDALQHAQCTLKSLVLRRCHFTSLSSEYLS 838

Qy       1002 LALMGNSWLTHLSLSMNPVEDNGVKLLCEVMREPSCHLQDLELVKCHLTAACCESLSCVI 1061
Db        839 TSLLHNKSLTHLDLGSNWLQDNGVKLLCDVFRHPSCNLQDLELMGCVLTNACCLDLASVI 898

Qy       1062 SRSRHLKSLDLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLTSDCCEALSLALSCN 1121
Db        899 LNNPNLRSLDLGNNDLQDDGVKILCDALRYPNCNIQRLGLEYCGLTSLCCQDLSSALICN 958

Qy       1122 RHLTSLNLVQNNFSPKGMMKLCSAFACPTSNLQIIGLWKWQ 1162
Db        959 KRLIKMNLTQNTLGYEGIVKLYKVLKSPKCKLQVLGQQDFQ 999
```

RESULT 3
US-10-511-989-14
; Sequence 14, Application US/10511989
; Publication No. US20060053496A1
; GENERAL INFORMATION:
; APPLICANT: University of North Carolina-Chapel Hill
; APPLICANT: Ting, Jenny
; APPLICANT: Linhoff, Michael
; APPLICANT: Harton, Johnathan
; APPLICANT: Williams, Kristi
; APPLICANT: Lich, John
; APPLICANT: O'Connor, William